JSPS Grants-in-Aid for Scientific Research (S)

Understanding Persistent Deflation in Japan

Working Paper Series

No. 040

April 2014

Zipf's Law, Pareto's Law, and the Evolution of Top Incomes in the U.S.

Shuhei Aoki

Makoto Nirei

UTokyo Price Project
702 Faculty of Economics, The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
Tel: +81-3-5841-5595
E-mail: watlab@e.u-tokyo.ac.jp
http://www.price.e.u-tokyo.ac.jp/english/

Zipf's Law, Pareto's Law, and the Evolution of Top Incomes in the U.S.*

Shuhei Aoki†
Faculty of Economics, Hitotsubashi University

Makoto Nirei‡
Institute of Innovation Research, Hitotsubashi University

April 8, 2014

Abstract

This paper presents a tractable dynamic general equilibrium model of income and firm-size distributions. The size and value of firms result from idiosyncratic, firm-level productivity shocks. CEOs can invest in their own firms' risky stocks or in risk-free assets, implying that the CEO's asset and income also depend on firm-level productivity shocks. We analytically show that this model generates the Pareto distribution of top income earners and Zipf's law of firms in the steady state. Using the model, we evaluate how changes in tax rates can account for the recent evolution of top incomes in the U.S. The model matches the decline in the Pareto exponent of income distribution and the trend of the top 1% income share in the U.S. in recent decades. In the model, the lower marginal income tax for CEOs strengthens their incentive to increase the share of their firms' risky stocks in their own asset portfolios. This leads to both higher dispersion and concentration of income in the top income group.

JEL Codes: D31, L11, O40
Keywords: income distribution; wealth distribution; Pareto exponent; top income share; firm size distribution; Zipf's law

*First draft: June 18, 2013. A previous version of this paper was entitled "Pareto Distributions and the Evolution of Top Incomes in the U.S." We thank Michio Suzuki, Tomoaki Yamada, Kozo Ueda, and Shin-ichi Fukuda for helpful comments and discussions. We acknowledge the financial support of a Grant-in-Aid for Scientific Research (KAKENHI 23243050) from JSPS.

†Address: 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan. Phone: +81 (42) 580-8477. Fax: +81 (42) 580-8195. Email: shuhei.aoki@gmail.com.

‡Address: 2-1 Naka, Kunitachi, Tokyo 186-8603, Japan. Phone: +81 (42) 580-8417. Fax: +81 (42) 580-8410. E-mail: nirei@iir.hit-u.ac.jp.

1 Introduction

For the last three decades, there has been a secular trend of concentration of income among the top earners in the U.S. economy. According to Alvaredo et al. (2013), the top 1% income share, the share of total income going to the richest top 1% of the population, declined from around 18% to 8% after the 1930s, but the trend was reversed during the 1970s. Since then, the income share of the top 1% has grown and had reached 18% by 2010, on par with the prewar level.

Along with the increasing trend in the top income share, a widening dispersion of income within the top income group has also been observed over the same periods. It is known that the tail part of the income distribution is described by a Pareto distribution very well. When income follows a Pareto distribution with exponent λ, the ratio of the number of people who earn more than x_1 to those who earn more than x_2, for any income levels x_1 and x_2, is $(x_1/x_2)^{-\lambda}$. Thus, the Pareto exponent λ is a measure of equality among the rich. The estimated Pareto exponent historically shows a close connection with the top income share. This exponent declined from 2.5 in 1975 to 1.6 in 2010, along with the secular increase in the top 1% income share.

There has been much debate about the causes of income concentration in recent decades. We pay special attention to the decrease in the marginal income tax rate as a driving force of income dispersion among the rich. The purpose of this paper is to develop a tractable dynamic general equilibrium model of income distribution, and then, use the model to analyze how a decrease in the marginal income tax rate affects income concentration.

Our main focus is income distribution; nevertheless, we require the model to be consistent with firm-side stylized facts. A substantial part of the concentration of income in recent decades is due to the increase in the incomes of top corporate executives and entrepreneurs (Piketty and Saez, 2003, Atkinson et al., 2011, and Bakija et al., 2012). The pay and assets of a CEO strongly depend on his firm's performance (see Frydman and Jenter, 2010 for a survey). In standard neoclassical models, a firm's performance is determined by its productivity. Therefore, a model of income concentration should be consistent with the stylized facts on the firm's productivity. Zipf's law is one of these facts. Zipf's law states that the firm size distribution, which is generated from the firm's productivity shocks in standard models (e.g., Luttmer, 2007), follows a special case of Pareto distribution with exponent $\lambda = 1$. Zipf's law is closely related to Gibrat's law, which refers to the fact that the growth rate of a firm is independent of its size (see Gabaix, 2009 and Luttmer, 2010).[1] We construct our model to be consistent with these laws.

We develop a model of heterogeneous firms and the CEOs' portfolio choices. In the model, the firms' size and value result from idiosyncratic, firm-level productivity shocks. CEOs can invest in their own firms' risky stocks or in risk-free assets. The dispersion of CEOs' incomes is determined by the risk taken in their after-tax portfolio returns.

The contribution of the paper is summarized as follows. First, this paper presents a parsimonious neoclassical growth model that generates Zipf's and Gibrat's laws of firms and Pareto's law of incomes from idiosyncratic, firm-level productivity shocks. The model is simple enough to allow analytical derivation of the stationary distributions of firms and income. Second, we obtain an analytical expression for the evolution of the probability density distribution of income in the transition path. Using this expression, we can numerically compute the transition dynamics of income distribution since an unanticipated and permanent cut in top marginal income tax rate. Third, we calibrate the model parameters and show that the transition path of the model computed as above matches the decline in the Pareto exponent of income distribution and the trend of increasing top income share observed in the last three decades. Hence, we argue that the calibrated analysis of our model predicts that the tax cut and CEOs' response to tax in their portfolio can explain the widening dispersion and higher concentration of income occurred in the U.S. The calibrated model also brings out testable implications for CEO portfolios and future development of inequality under the current tax rate level.

Piketty and Saez (2003) argue that a cut in the top marginal income tax rate is a plausible reason for the recent evolution of top incomes, as compared with other reasons such as skill-biased technical change. Piketty et al. (2011) report that among the OECD countries, the countries that have experienced a sharp rise in the top 1% income share are also the ones where the top marginal income tax rate has reduced drastically. Our paper shares their view that a tax cut is an important factor. However, our model differs from theirs in that a cut in top marginal income tax rate itself does not matter, as in the case that dividend tax in the "new" explanation of dividend taxation (Sinn, 1991 and McGrattan and Prescott, 2005) does not affect investment decisions. Instead, in our model, a cut in top marginal income tax rate relative to other taxes, such as capital gains and corporate taxes, does affect CEOs' portfolio choices and

[1] Note that as Gabaix (2009) and Luttmer (2010) point out, deviations from Gibrat's law are reported for young and small firms. However, we exclude these issues from our analysis, because our focus is on the evolution of top earners who manage big firms.

the wealth and income distributions.

Recently, several papers have built models to understand why income distribution follows a Pareto distribution. There are two types of approaches in the literature. The first explains Pareto's law of incomes by the assumption that other variables follow certain types of distributions. Gabaix and Landier (2008) take this approach. They construct a model of the CEO's pay that assumes that the firm size distribution follows Zipf's law and the CEO's talent follows a certain distribution. Under the settings, they show that the CEO's pay follows a Pareto distribution. Their model has the advantage of being consistent with the two stylized facts, that is, Zipf's law of firms and Pareto's law of incomes. However, their model deals with the case where the Pareto exponent is constant. Jones and Kim (2012) extend the model to be consistent with the recent decline in the Pareto exponent of income distribution in the U.S. As compared to the papers taking this approach, our paper's contribution is to build a model that generates Zipf's and Pareto's laws, both from the productivity shocks of firms, without assuming particular types of distributions.

The second approach explains Pareto's law of incomes by idiosyncratic shocks. Using a household model with a consumption function, Nirei and Souma (2007) show that idiosyncratic shocks on the household's asset returns generate Pareto's law of assets and incomes. Benhabib et al. (2011 and 2012) show a similar result for a model of households that optimally make saving and bequest decisions. These models are not dynamic general equilibrium models, in the sense that they only consider the household's problem and not the firm's. Nirei (2009) extends the framework to a Bewley-type model and shows that idiosyncratic shocks on firms' productivities generate Pareto's law of incomes in a dynamic general equilibrium environment. Toda (2012) also builds a similar, but more analytically tractable, dynamic general equilibrium model and derive Pareto's law. Our study follows this approach.[2] As compared with previous studies, this paper features a model that can explain Zipf's law of firms, and analyzes how the recent tax cut affects the evolution of top incomes.[3]

Perhaps, the closest paper to ours is Kim (2013), who follows the latter approach and builds a model of human capital accumulation with idiosyncratic shocks that generates Pareto's law of incomes (see

[2]Different from Benhabib et al. (2011), who adopt the overlapping generations setting, our model, adopting the perpetual youth setting, does not take into account the bequest motive of households. Its justification comes from a finding in Kaplan and Rauh (2013), who report, "Those in the Forbes 400 are less likely to have inherited their wealth or to have grown up wealthy."

[3]This paper's model is also consistent with the fact that the firm's productivity distribution also follows a Pareto distribution (Mizuno et al., 2012).

also Jones and Kim, 2013, who incorporate creative destruction into the model). Using the model, she analyzes the impact of a cut in top marginal income tax in recent decades on the Pareto exponent of income distribution. As compared with her paper, our paper's contribution is to build a model that also explains Zipf's law of firms, from the same shocks that generate Pareto's law of incomes. In addition, because the mechanism through which a tax cut affects top incomes is different from hers, the predictions of the models are also different. For example, in her model, an income tax cut encourages human capital accumulation among top income earners. This would result in the increase in the level of the per capita output in the U.S. in recent decades, as compared with previous periods and other countries such as France. In contrast, in our model, a tax cut does not directly affect capital accumulation.

Finally, our model is also closely related with the general equilibrium models of firm size distribution that explain Zipf's law of firms (for a survey, see Luttmer, 2010). The basic mechanism employed in our study to generate Zipf's law of firms draws on the literature. In comparison to the literature, our firm-side formulation is rather simplified, because our focus is to understand the evolution of top incomes.

The organization of the paper is follows. Section 2 sets up a dynamic general equilibrium model. Section 3 discusses the firm-side properties of the model and derives Zipf's law of firms. Section 4 analyzes the aggregate dynamics of the model and defines the equilibrium. After defining the equilibrium, Section 5 illustrates how in the steady state, the household asset and income distribution follows a Pareto distribution. Section 6 analyzes how a tax cut affects top incomes in our model and contrasts the results with the data. Finally, in Section 7, we present our concluding remarks.

2 Model

It is well known that the stationary distribution of certain types of stochastic processes follows a Pareto distribution. The purpose of the model presented here is to incorporate these stochastic processes into an otherwise standard general equilibrium model with incomplete markets and replicate Pareto distributions observed as stylized facts. Key assumptions that generate Zipf's law of firms are that the firm's productivity is affected by multiplicative idiosyncratic shocks and there is a lower bound for the firm size. Similarly, key assumptions that generate Pareto's law of the households' assets and incomes are that these assets are affected by multiplicative idiosyncratic shocks and each household faces a constant probability of death (that is, the perpetual youth assumption). In the next sections, we discuss how these properties generate the laws.

2.1 Households

There is a continuum of households with a mass L. As in Blanchard (1985), by a Poisson hazard rate ν, each household is discontinued and is replaced by a newborn household that has no bequest. Households participate in a pension program. If a household dies, all of his non-human assets are distributed to living households. The amount a living household gets is the pension premium rate ν times his financial assets.

The households consist of entrepreneurs and workers. A mass N of households are entrepreneurs and the remaining $L - N$ are workers. Each provides one unit of labor and earns wage income w_t. They also get government transfer tr_t. Among these households, only entrepreneurs manage firms. An entrepreneur managing his firm has the benefit of holding the stocks of his firm relatively cheaper, as is explained shortly. Entrepreneurs leave their firm and become workers with a Poisson hazard rate p_f. Thus, there are two types of workers, namely, workers who were previously entrepreneurs and workers by birth. We refer to the former as former entrepreneurs and the latter as innate workers.[4]

These households maximize expected discounted log utility

$$\mathsf{E}_t \int_t^\infty \ln c_{i,s} e^{-(\beta+\nu)s} ds,$$

where β is the discount rate, by optimally choosing sequences of consumption $c_{i,s}$ and an asset portfolio. As the asset portfolio, a worker can hold (i) a risk-free market portfolio $b_{i,t}$ that consists of the market portfolio of firms' stocks, and (ii) human assets h_t that consist of wage incomes w_t and government transfers tr_t. The risk-free market portfolio yields a net return r_t^f (and pension premium ν) with certainty. The human asset is defined by $h_t = \int_t^\infty (w_u + tr_u) e^{-\int_t^u (\nu + r_s^f) ds} du$, whose return is

$$(\nu + r_t^f) h_t = (w_t + tr_t) + dh_t/dt.$$

An entrepreneur can hold (i) a risk-free market portfolio $b_{i,t}$ and (ii) human assets h_t, similar to a worker. In addition, the entrepreneur can also hold (iii) risky stocks of his firm $s_{i,t}$. Owing to the setup

[4] We introduce the former entrepreneurs for a purely quantitative reason. The qualitative results of this study are intact even when $p_f = 0$. Quantitatively, if we do not introduce former entrepreneurs and all of the entrepreneurs retain their positions, the mobility of a household's asset or income level becomes too slow or the Pareto exponent of income distribution becomes too low, as compared with the data.

of the model described in the following sections, the risky stocks are affected by uninsurable idiosyncratic shocks; the risk and returns of an entrepreneur's risky stocks are ex ante identical across entrepreneurs; and the expected return of risky stocks exceeds that of the risk-free market portfolio because transaction costs and tax rates differ between the two assets. Let $q_{i,t}$ and $d_{i,t}$ be the price and dividend of the risky stocks, respectively. Then, the return of the risky stock is described by the following stochastic process:

$$((1-\tau^e)d_{i,t}dt + dq_{i,t})/q_{i,t} = \mu_{q,t}dt + \sigma_{q,t}dB_{i,t},$$

where τ^e is the tax rate on the risky stock and $B_{i,t}$ is a Wiener process. Note that we interpret holding risky stocks in the model as a CEO's incentive scheme in the real world. In the numerical analysis, we calibrate tax on risky stocks, τ^e, by the top marginal tax rate on ordinary income imposed on the CEO's pay. We discuss the similarity of our formulation with previous studies on CEO pay and compare our model's prediction with the data in Section 6.5.1.

Let $a_{i,t} = s_{i,t}q_{i,t} + b_{i,t} + h_t$ denote the total wealth of a household. (Note that if the household is a worker, $s_{i,t} = 0$.) The total wealth accumulates according to the following process:

$$\begin{aligned} da_{i,t} &= (\nu(s_{i,t}q_{i,t} + b_{i,t}) + \mu_{q,t}s_{i,t}q_{i,t} + r_t^f b_{i,t} + (\nu + r_t^f)h_t - c_{i,t})dt \\ &\quad + \sigma_{q,t}s_{i,t}q_{i,t}dB_{i,t} \\ &= \mu_{a,t}a_{i,t}dt + \sigma_{a,t}a_{i,t}dB_{i,t}. \end{aligned} \tag{1}$$

where $\mu_{a,t}a_{i,t} \equiv \nu a_{i,t} + \mu_{q,t}x_{i,t}a_{i,t} + r_t^f(1-x_{i,t})a_{i,t} - c_{i,t}$, $\sigma_{a,t}a_{i,t} \equiv \sigma_{q,t}x_{i,t}a_{i,t}$, and $x_{i,t}$ is the share of $a_{i,t}$ invested in the risky stocks. $dB_{i,t}$ is a multiplicative shock to the asset accumulation, in that the shock is multiplied by the current asset level $a_{i,t}$.

The household's dynamic programming problem is specified as follows:

$$V^i(a_{i,t}, \mathbf{S}_t) = \max_{c_{i,t}, x_{i,t}} \ln c_{i,t}dt + e^{-(\beta+\nu)dt}\,\mathsf{E}_t[V^{i'}(a_{i,t+dt}, \mathbf{S}_{t+dt})] \tag{2}$$

subject to (1), where $\mathbf{S}_t$ is a set of variables that describes the aggregate dynamics of the model (for the definition, see Section 4) and V^i denotes value functions of household characteristics i. That is, if the household is an entrepreneur, $i = e$; and if he is a worker, $i = \ell$ (more specifically, if he is an innate worker, $i = w$; and if he is a former entrepreneur, $i = f$). Note that if the household is an entrepreneur, the household characteristics in the next period, denoted by i', can be both entrepreneur and worker. If the household is an innate worker or a former entrepreneur, $i' = i$.

The household problem is a variant of Merton's dynamic portfolio problem (Merton, 1969, 1971, 1973, Campbell and Viceira, 2002, and Benhabib et al., 2012). It is well known that the solution of the problem under the log utility follows the myopic rules

$$x_{i,t} = \begin{cases} \frac{\mu_{q,t} - r_t^f}{\sigma_{q,t}^2}, & \text{if } i = e, \\ 0, & \text{otherwise,} \end{cases} \tag{3}$$

$$v_{i,t} = \beta + \nu, \tag{4}$$

where $v_{i,t}$ is the consumption–wealth ratio (see Appendix A for derivations), and satisfies the transversality condition

$$\lim_{T \to \infty} e^{-(\beta+\nu)T} \mathsf{E}_0 \left[V^i(a_{i,T}, \mathbf{S}_T) \right] = 0. \tag{5}$$

In the model, we assume that entrepreneurs can hold risky stocks of their own firms. We can relax the assumption and allow households to hold risky stocks of the firms not managed by the households, whose expected returns are as low as that of risk-free assets, r^f, owing to transaction costs and a different tax rate that are explained in the next section. Then, because the shocks on risky stocks are assumed to be uncorrelated with one another, the optimal portfolio share of another firm's risky stocks $x'_{i,t}$ becomes $(r_t^f - r_t^f)/(\sigma'_{q,t})^2 = 0$, where $\sigma'_{q,t}$ is the volatility of these risky stocks (see e.g., Campbell and Viceira, 2002). This implies that the results are unchanged even when the assumption is relaxed.

2.2 Firms and the financial market

A continuum of firms with a mass N produces differentiated goods. As in McGrattan and Prescott (2005), each firm issues shares and owns and self-finances capital $k_{j,t}$. As noted above, the entrepreneur of the firm can directly own shares of his firm. Financial intermediaries also own the firm's shares, and by combining the shares of all of these firms, issue risk-free market portfolios to households. This helps to diversify the idiosyncratic shocks of the firms. The financial intermediaries incur ι per dividend $d_{j,t}$

as transaction costs. We assume that financial intermediaries own the majority shares, or that when an entrepreneur owns his firm's shares in the form of preferred stocks without voting rights. Under the setup, firms maximize expected profits following the interest of financial intermediaries. Then, the market value of a firm becomes the net present value of the after-tax profits discounted by the risk-free rate r_t^f. We make these assumptions to simplify the analysis.

2.2.1 Financial intermediary's problem

In this model, returns and risks on risky stocks are ex ante identical across firms and that shocks on the risky stocks are uncorrelated with each other. Then, a financial intermediary maximizes residual profit by diversifying the risks on risky stocks and issuing risk-free assets as follows:

$$\max_{s_{j,t}^f} \mathsf{E}_t \left[\int_0^N \left\{ (1-\tau^f-\iota) d_{j,t} dt + dq_{j,t} \right\} s_{j,t}^f dj \right] - r_t^f dt \left(\int_0^N q_{j,t} s_{j,t}^f dj \right),$$

where $s_{j,t}^f$ is the shares of firm j owned by the financial intermediary and τ^f is the dividend tax, which is different from the tax rate on risky stocks τ^e. We interpret τ^f in the numerical analysis as a combination of capital gains and corporate income taxes. In Section 6, we account for the evolution of top incomes by the change in the difference between these tax rates. The solution of the problem leads to

$$r_t^f q_{j,t} dt = \mathsf{E}_t[(1-\tau^f-\iota) d_{j,t} dt + dq_{j,t}]. \tag{6}$$

2.2.2 Firm's problem

There are heterogeneous firms in the economy. The production function of firm j is

$$y_{j,t} = z_{j,t} k_{j,t}^{\alpha} \ell_{j,t}^{1-\alpha}.$$

The productivity of the firm evolves as

$$dz_{j,t} = \mu_z z_{j,t} dt + \sigma_z z_{j,t} dB_{j,t},$$

where $B_{j,t}$ is a Wiener process that is uncorrelated with shocks in other firms. $dB_{j,t}$ is a multiplicative shock to the productivity growth, because the shock is multiplied by its productivity level $z_{j,t}$. Under the formulation, when the firm's size is proportional to its productivity, as will be shown below, Gibrat's law of firms holds; that is, the growth rate of the firm is independent of the firm's size.

In order to derive the property that the firm-size distribution follows Zipf's law, we impose the following assumptions on the minimum level of firm size. Following Rossi-Hansberg and Wright (2007), who construct a model of establishment size dynamics, we assume that there is a minimum level of employment $\ell_{\min}$, that is,

$$\ell_{j,t} \geq \ell_{\min}.$$

A firm whose optimal employment is less than $\ell_{\min}$ is restructured. More precisely, we define the productivity level $z_{\min}$ as the one at which, when the firm optimally chooses labor (following (7) below), $\ell_{j,t} = \ell_{\min}$. We assume that the firm whose productivity $z_{j,t}$ is less than $z_{\min}$ has to be restructured such that the firm buys productivities and accompanying capital from other firms at the market price, in order to increase its own size. Correspondingly, we assume that each firm sells a constant fraction of its capital to the firms undergoing restructuring (in the next section, we discuss how these deals are conducted).

A firm chooses the investment level $dk_{j,t}$ and employment $\ell_{j,t}$ to maximize profit as follows:

$$r_t^f q(k_{j,t}, z_{j,t}, \mathbf{S}_t)dt = \mathsf{E}_t \left[\max_{dk_{j,t}, \ell_{j,t}} (1 - \tau^f - \iota) d_{j,t} dt + dq(k_{j,t}, z_{j,t}, \mathbf{S}_t) \right]. \tag{7}$$

The dividend $d_{j,t}$ consists of

$$d_{j,t}dt = (p_{j,t}y_{j,t} - w_t \ell_{j,t} - \delta k_{j,t})\, dt - dk_{j,t},$$

where $p_{j,t}$ and $y_{j,t}$ are, respectively, the price and quantity of the good produced by the firm, $k_{j,t}$ is the capital, w_t is the wage rate, and δ is the depreciation rate.

By solving the firm's maximization problem, we obtain the following conditions (see Appendix B for details):

$$\mathsf{MPK}_t \equiv r_t^f + \delta = \frac{\partial p_{j,t} y_{j,t}}{\partial k_{j,t}}, \tag{8}$$

$$w_t = \frac{\partial p_{j,t} y_{j,t}}{\partial \ell_{j,t}}. \tag{9}$$

Two remarks need to be made about the firm's problem. First, in the model, the marginal product of capital, MPK, becomes the same among firms, because the stochastic discount factor of those who own diversified bonds is not correlated with the shock of firm j. Second, because taxes in the model are imposed on dividends, as found in the "new view" literature of dividend taxation (Sinn, 1991 and McGrattan and Prescott, 2005), they do not distort MPK.

2.3 Aggregation and market conditions

We now consider the market conditions for the aggregate economy. (Throughout the paper, we use upper case letters to denote aggregate variables.) Goods that a mass N of firms produces are aggregated according to

$$Y_t = \left(\int_0^N \left(\frac{1}{N} \right)^{1-\rho} y_{j,t}^{\rho} dj \right)^{\frac{1}{\rho}}. \tag{10}$$

We assume that the aggregate good Y is produced competitively. Other aggregate variables are simply summed up over the households or firms. For example, $C_t = \int_0^L c_{i,t} di$ and $K_t = \int_0^N k_{j,t} dj$.

The market clearing condition for final goods is

$$C_t + \frac{dK_t}{dt} - \delta K_t + \iota \left(1 - \frac{A_{e,t} x_{e,t}}{Q_t} \right) D_t = Y_t,$$

where $A_{e,t}$ is the total assets of entrepreneurs and Q_t is the aggregate financial asset, the sum of risk-free market portfolios and risky stocks. The last term on the left-hand side of the equation indicates that a part of the final goods is consumed as transaction costs. The labor market clearing condition is

$$\int_0^N \ell_{j,t} dj = L. \tag{11}$$

The market clearing condition for the shares of firms is

$$s_{j,t} + s^f_{j,t} = 1,$$

where $s_{j,t}$ is the shares owned by the entrepreneur according to (3) and $s^f_{j,t}$ is the shares owned by financial intermediaries. We assume that government transfers are adjusted such that tax revenues equal government transfers in each period.

3 Firm-Side Properties

Before we define the equilibrium and solve the model, we review some of its firm-side properties. First, in this model, given r^f_t, the firm-side variables, such as $\ell_{j,t}$, $k_{j,t}$, and $d_{j,t}$ can be obtained as closed-form expressions. These variables can be written as a product of components that are common across firms and the heterogeneous component. Second, the firm's productivity distribution is obtained independently of other variables. It is a Pareto distribution that establishes Zipf's law of firms when the minimum employment level $\ell_{\min}$ is sufficiently small.

3.1 Firm-side variables

Employing the firm's first-order conditions (FOCs), (8) and (9), together with the aggregate condition (10) and the labor market condition (11), the firm's variables can be written as follows (for the derivations, see Appendices B.2 and B.3):

$$\ell_{j,t} = \overline{\ell}_t \; z_{j,t}^{\phi}, \text{ where } \overline{\ell}_t \equiv \left(\frac{L/N}{\overline{\mathsf{E}}\left\{ z_{j,t}^{\phi} \right\}} \right) \text{ and } \phi \equiv \frac{\rho}{1-\rho}, \tag{12}$$

$$p_{j,t}y_{j,t} = \overline{py}_t \overline{\ell}_t \; z_{j,t}^{\phi}, \text{ where } \overline{py}_t \equiv \left(\frac{\alpha\rho}{\mathsf{MPK}_t} \right)^{\frac{\alpha}{1-\alpha}} \overline{\mathsf{E}}\left\{ z_{j,t}^{\phi} \right\}^{\frac{1}{\phi}\frac{1}{1-\alpha}}, \tag{13}$$

$$k_{j,t} = \overline{k}_t \overline{\ell}_t z_{j,t}^{\phi}, \text{ where } \overline{k}_t \equiv \left(\frac{\alpha\rho}{\mathsf{MPK}_t} \right)^{\frac{1}{1-\alpha}} \overline{\mathsf{E}}\left\{ z_{j,t}^{\phi} \right\}^{\frac{1}{\phi}\frac{1}{1-\alpha}}, \tag{14}$$

$$d_{j,t}dt = \overline{d}_t \overline{\ell}_t z_{j,t}^{\phi} dt - \phi\sigma_z \overline{k}_t \overline{\ell}_t z_{j,t}^{\phi} dB_{j,t},$$

$$\text{where } \overline{d}_t \equiv (1-(1-\alpha)\rho)\overline{py}_t - (\delta + \mu_{k,t})\,\overline{k}_t,$$

$$q_{j,t} = \overline{q}_t \overline{\ell}_t z_{j,t}^{\phi}, \quad \text{where } \overline{q}_t \equiv (1-\tau^f-\iota)\overline{d}_t \int_t^{\infty} \exp\left\{ - \int_t^u (r_s^f - \mu_{d,s}) ds \right\} du. \tag{15}$$

Note that $\overline{\mathsf{E}}\left\{ z_{j,t}^{\phi} \right\}$ is the average of $z_{j,t}^{\phi}$ over the firm size distribution (we will show later that the average exists), and $\mu_{k,t}$ and $\mu_{d,t}$ are the expected growth rates of $k_{j,t}$ and $d_{j,t}$, respectively.

In the above equations, each of the variables has common components, such as $\overline{\ell}_t$ and $\overline{py}_t$ and the heterogeneous component, $z_{j,t}^{\phi}$. Thus, the size distributions of the firm-side variables depend only on the heterogeneous component.

3.2 Restructuring

Before deriving the firm size distribution, we analyze how firms that are restructuring buy the assets of other firms. In each small time interval, some firms decrease their productivities from t and $t+dt$ to $z_{j,t+dt} < z_{\min}$. We assume that these firms increase the productivities and firm sizes by mergers and acquisitions (M&A) through the stock market.

To understand how the productivities of firms change through M&A, suppose that the buyer firm j, whose firm value is $q_{j,t+dt}^{\text{before}}$, acquires a part of the seller firm by paying $q_{j,t+dt}^{\text{M\&A}}$. The firm value after the M&A is $q_{j,t+dt}^{\text{after}} = q_{j,t+dt}^{\text{before}} + q_{j,t+dt}^{\text{M\&A}}$. We assume that according to (15), the productivity of the buyer firm after M&A $z_{j,t+dt}^{\text{after}}$ increases, thereby satisfying $q_{j,t+dt}^{\text{after}} = \overline{q}_{t+dt}\overline{\ell}_{t+dt}(z_{i,t+dt}^{\text{after}})^{\phi}$. The productivity of the seller firm decreases by $q_{j,t+dt}^{\text{M\&A}}$, also according to (15).

Thus, on the buyer side, firms whose productivities at $t+dt$ are less than $z_{\min}$ have to buy $\overline{q}_{t+dt}\overline{\ell}_{t+dt}\left(z_{\min}^{\phi} - z_{j,t+dt}^{\phi}\right)$ from other firms to undergo restructuring. We denote the total of these payments by the restructured firms as $Q_{\text{restructuring},t+dt}$.

On the seller side, we assume that at each instant, every firms sell a constant fraction $m\phi dt$ of its value to these restructuring firms at the market price $m\phi q_{j,t}dt$. (ϕ is multiplied as the adjustment term.) The total value of the sellouts is

$$m\phi dt \int_0^N q_{j,t}dj = N\overline{q}_{t+dt}\overline{\ell}_{t+dt}\overline{\mathsf{E}}\left\{z_{j,t}^{\phi}\right\} m\phi dt.$$

Because the demand of restructuring firms equates the supply,

$$Q_{\text{restructuring},t+dt} = N\overline{q}_{t+dt}\overline{\ell}_{t+dt}\overline{\mathsf{E}}\left\{z_{j,t}^{\phi}\right\} m\phi dt. \tag{16}$$

Rearranging this equation and taking the limit as dt approaches zero from above, we obtain (see Appendix B.4 for details)

$$m = (\lambda - \phi)\frac{\sigma_z^2}{4}, \tag{17}$$

where λ is the Pareto exponent of the firm size distribution that is pinned down in the next section.

3.3 Firm size distribution

We detrend the firm's productivity to derive the invariant productivity distribution. Let $\tilde{z}_{j,t}$ be the firm's productivity level after selling a part of the firm's assets to restructuring firms, detrended by $e^{g_z t}$ (g_z is a constant whose value is determined below). The firm's detrended productivity growth after a sellout is

$$\begin{aligned} d\tilde{z}_{j,t} &= (\mu_z - g_z - m)\,\tilde{z}_{j,t}dt + \sigma_z\tilde{z}_{j,t}dB_{j,t}, \\ \text{or, } d\ln\tilde{z}_{j,t} &= \left(\mu_z - g_z - \frac{\sigma_z^2}{2} - m\right)dt + \sigma_z dB_{j,t}. \end{aligned} \tag{18}$$

The Fokker–Planck equation for the probability density $f_z(\ln\tilde{z}_{j,t}, t)$ for the firm's productivity is

$$\frac{\partial f_z(\ln\tilde{z}_{j,t}, t)}{\partial t} = -\left(\mu_z - g_z - \frac{\sigma_z^2}{2} - m\right)\frac{\partial f_z(\ln\tilde{z}_{j,t}, t)}{\partial\ln\tilde{z}_{j,t}} + \frac{\sigma_z^2}{2}\frac{\partial^2 f_z(\ln\tilde{z}_{j,t}, t)}{\partial(\ln\tilde{z}_{j,t})^2}.$$

In this paper, we assume an invariant distribution for firms, that is, $\partial f_z(\ln\tilde{z}_{j,t}, t)/\partial t = 0$. In the case of the invariant distribution, the Fokker–Planck equation has a solution in exponential form,

$$f_z(\ln \tilde{z}_{j,t}) = F_0 \exp(-\lambda \ln \tilde{z}_{j,t}), \tag{19}$$

where the coefficients satisfy

$$F_0 = \lambda \tilde{z}_{\min}^{\lambda}, \quad \lambda = -2\left(\mu_z - g_z - \frac{\sigma_z^2}{2} - m\right)/\sigma_z^2. \tag{20}$$

(19) shows that the distribution of $\ln \tilde{z}_{j,t}$ follows an exponential distribution. Through a change of variables, it can also be shown that the distribution of $\tilde{z}_{j,t}$ follows a Pareto distribution whose Pareto exponent is λ.

In this model, the exogenous parameter $\ell_{\min}$ pins down λ and g_z. From the restriction on $\ell_{\min}$ and (12) (and by employing (22) below), we obtain the Pareto exponent for $\tilde{z}_{j,t}$ as

$$\lambda = \frac{1}{1 - \frac{\ell_{\min}}{L/N}}\phi. \tag{21}$$

With this λ, we obtain the rescaling parameter g_z that assures the existence of the invariant distribution of $\tilde{z}_{j,t}$ from (17) and (20).

Four remarks need to be made on the firm size distribution. First, we obtain a constant rescaled mean $\overline{\mathsf{E}}\left\{\tilde{z}_{j,t}^{\phi}\right\}$ for a constant $\tilde{z}_{\min}$ as follows:

$$\overline{\mathsf{E}}\left\{\tilde{z}_{j,t}^{\phi}\right\} = \int_{\tilde{z}_{\min}}^{\infty} \tilde{z}^{\phi} f_z(\ln \tilde{z}) \frac{\partial \ln \tilde{z}}{\partial \tilde{z}} d\tilde{z} = \frac{F_0 \tilde{z}_{\min}^{-(\lambda-\phi)}}{\lambda - \phi}. \tag{22}$$

Second, the growth rate of the aggregate output is $g \equiv g_z/(1-\alpha)$. We can confirm this property by detrending and aggregating (13).

Third, Zipf's law holds for the distribution of firm size, $\ell_{j,t}$ and $y_{j,t}$. This is because the firm size distribution cross-sectionally obeys $\tilde{z}_{j,t}^{\phi}$, whose Pareto exponent is λ/ϕ. (21) shows that $\lambda/\phi > 1$ and that if $\ell_{\min}$ is sufficiently small as compared with the average employment level L/N, λ/ϕ becomes close to 1.

Fourth, the expected growth rate of the detrended firm side variables, that is, the expected growth rate of $\tilde{z}_{j,t}^{\phi}$, is negative when the firm does not restructure. We can show this property, that is,

$\phi\left(\mu_z - g_z - m + (\phi-1)\frac{\sigma_z^2}{2}\right) < 0$, from (20), if $\lambda/\phi \geq 1$, which is satisfied when the third remark holds. This is a key property that generates a Pareto distribution with a finite distributional mean, because otherwise, the distribution diffuses over time.

4 Aggregate Dynamics and Equilibrium of the Model

In this model, because the household's policy functions are independent of its wealth level, the dynamics of aggregate variables are obtained independent of the heterogeneity within entrepreneurs, innate workers, and former entrepreneurs. In this section, we first show this property and then, define the equilibrium of the model.

4.1 Aggregate dynamics of the model

Let variables with tilde such as $\tilde{K}_t$ be the variables detrended by e^{gt}. We show below that the aggregate dynamics of the detrended variables can be reduced to the differential equations of $\tilde{\mathbf{S}}_t \equiv \mathbf{S}_t/e^{gt} = (\tilde{A}_{e,t}, \tilde{A}_{w,t}, \tilde{A}_{f,t}, \tilde{H}_t, \tilde{K}_t)$, i.e.,

$$d\tilde{\mathbf{S}}_t = \mu_{\tilde{\mathbf{S}}}(\tilde{\mathbf{S}}_t)dt. \tag{23}$$

Note that $\tilde{A}_{e,t}$, $\tilde{A}_{w,t}$, and $\tilde{A}_{f,t}$ are the (detrended) aggregate productivities of entrepreneurs, innate workers, and former entrepreneurs. We also show that price variables, $\tilde{\mathbf{r}}_t \equiv (r_t^f, \mu_{q,t}, \sigma_{q,t})$, are the functions of $\tilde{\mathbf{S}}_t$, i.e.,

$$\tilde{\mathbf{r}}_t = f_{\tilde{\mathbf{r}}}(\tilde{\mathbf{S}}_t). \tag{24}$$

These results are obtained by the following steps.

1. Given $\tilde{K}_t$, from (14),

$$r_t^f + \delta = \mathsf{MPK}_t = \alpha\rho\overline{\mathsf{E}}\left\{\tilde{z}_{j,t}^{\phi}\right\}^{\frac{1}{\phi}} / \left(\frac{\tilde{K}_t}{L}\right)^{1-\alpha}.$$

2. From (4), $\tilde{C}_t = (\beta+\nu)\tilde{A}_t$. $\tilde{Q}_t = \tilde{A}_t - \tilde{H}_t$. Given MPK_t, $\tilde{Y}_t = \overline{py}_t L/e^{gt}$ is pinned down. Then,

$$\frac{d\tilde{K}_t}{dt} = \tilde{Y}_t - \delta\tilde{K}_t - \tilde{C}_t - \iota\left(1 - \frac{\tilde{A}_{e,t}x_t}{\tilde{Q}_t}\right)\tilde{D}_t - g\tilde{K}_t,$$
$$\tilde{D}_t = (1-(1-\alpha)\rho)\tilde{Y}_t - (\delta+\mu_{\tilde{K}})\tilde{K}_t - \frac{d\tilde{K}_t}{dt}, \text{ where } \mu_{\tilde{K}} \equiv g + \phi\left\{(\mu_z - g_z) + (\phi-1)\frac{\sigma_z^2}{2}\right\},$$

and $x_{e,t}$ are jointly determined (see Appendix B.2 for the derivation of the latter equation).

Note that here, the expected return and volatility of a risky stock are jointly determined as follows (see Appendix B.3 for details of the derivations):

$$\mu_{q,t} = \left\{\left(\frac{1-\tau^e}{(1-\tau^f-\iota)} - 1\right)\frac{1}{\int_t^\infty \exp\left\{-\int_t^u (r_s^f - \mu_{d,s})ds\right\}du} + r_t^f\right\},$$
$$\sigma_{q,t} = \phi\sigma_z \times \left\{1 - \left(\frac{1-\tau^e}{(1-\tau^f-\iota)}\right)\frac{\tilde{K}_t}{\tilde{D}_t}\frac{1}{\int_t^\infty \exp\left\{-\int_t^u (r_s^f - \mu_{d,s})ds\right\}du}\right\},$$

where $\int_t^\infty \exp\left\{-\int_t^u (r_s^f - \mu_{d,s})ds\right\}du$ is computed by the following equation:

$$\int_t^\infty \exp\left\{-\int_t^u (r_s^f - \mu_{d,s})ds\right\}du = \frac{\tilde{Q}_t}{(1-\tau^f-\iota)\tilde{D}_t}.$$

3. The assets for the three types of households evolve as follows:

$$\frac{d\tilde{A}_{e,t}}{dt} = (\mu_{ae,t} - g)\,\tilde{A}_{e,t} + (\nu+p_f)N\tilde{H}_t/L - (\nu+p_f)\tilde{A}_{e,t},$$
$$\frac{d\tilde{A}_{w,t}}{dt} = (\mu_{a\ell,t} - g)\,\tilde{A}_{w,t} + (\nu L - (\nu+p_f)N)\tilde{H}_t/L - \nu\tilde{A}_{w,t},$$
$$\frac{d\tilde{A}_{f,t}}{dt} = (\mu_{a\ell,t} - g)\,\tilde{A}_{f,t} + p_f\tilde{A}_{e,t} - \nu\tilde{A}_{f,t},$$

where $\mu_{ae,t}$ and $\mu_{a\ell,t}$ are the $\mu_{a,t}$s of an entrepreneur and a worker. The human asset evolves as

$$\frac{d\tilde{H}_t}{dt} = -(\tilde{w}_t + \tilde{tr}_t)L + (\nu + r_t^f - g)\tilde{H}_t, \tag{25}$$

where

$$\tilde{w}_t = (1-\alpha)\rho\tilde{Y}_t/L,$$
$$\tilde{tr}_t = \left\{ \frac{\tilde{A}_{e,t}x_{e,t}}{\tilde{Q}_t}\tau^e + \left(1 - \frac{\tilde{A}_{e,t}x_{e,t}}{\tilde{Q}_t}\right)\tau^f \right\} \tilde{D}_t/L.$$

4.2 Definition of a competitive equilibrium

Using the property on the aggregate dynamics, we now define the equilibrium of the model. To define the equilibrium, we specify the initial endowments of physical capitals and stocks in the following way. First, to simplify the analysis, in what follows, we focus on the equilibrium under which the initial capital of a firm is proportional to the firm's productivity, i.e., $\tilde{k}_{j,0} \propto \tilde{z}_{j,0}^{\phi}$. Then, the initial value of a firm is also proportional to the firm's productivity, i.e.,

$$\tilde{q}_{j,0} = \frac{\tilde{z}_{j,0}^{\phi}}{\bar{\mathsf{E}}\left\{\tilde{z}_{j,0}^{\phi}\right\}}\tilde{Q}_0, \text{ where } \tilde{Q}_0 = \tilde{A}_0 - \tilde{H}_0. \tag{26}$$

Second, we assume that all stocks are initially owned by households and except for those held by entrepreneurs are sold to the financial intermediaries at period 0.[5] Let $s_{j,0}^i$ be the initial shares of firm j held by household i (then, e.g., $\int_0^L s_{j,0}^i di = 1$).

A competitive equilibrium of the model, given the set of the firm's productivities, $\{\tilde{z}_{j,t}\}_{j,t}$, the initial capitals of firms, $\tilde{k}_{j,0} \propto \tilde{z}_{j,0}^{\phi}$, the initial shares of firms held by households, $\{s_{j,0}^i\}_{i,j}$, is a set of household variables, $\{x_{i,t}, v_{i,t}, \tilde{a}_{i,t}\}_{i,t}$, price variables, $\tilde{q}_{j,0}$ and $\{\tilde{\mathbf{r}}_t\}_t \equiv \{r_t^f, \mu_{q,t}, \sigma_{q,t}\}_t$, and aggregate variables, $\{\tilde{\mathbf{S}}_t\}_t \equiv \{\mathbf{S}_t/e^{gt}\}_t = \{\tilde{A}_{e,t}, \tilde{A}_{w,t}, \tilde{A}_{f,t}, \tilde{H}_t, \tilde{K}_t\}_t$ such that

- the household variables, $\{x_{i,t}, v_{i,t}, \tilde{a}_{i,t}\}_{i,t}$, where $\tilde{a}_{i,0} = \int_0^N \tilde{q}_{j,0}s_{j,0}^i dj + \tilde{H}_0/L$, are chosen according to the household's decisions on the portfolio choice (3) and (4), and the law of motion for total asset (1), and satisfy the transversality condition (5),
- the price variables, $\tilde{q}_{j,0}$ and $\{\tilde{\mathbf{r}}_t\}_t$, are determined by the aggregate variables $\tilde{\mathbf{S}}_t$ according to (26) and (24),
- and the aggregate variables, $\{\tilde{\mathbf{S}}_t\}_t$, evolve according to (23).

[5] We assume that the sellout to the financial intermediaries is mandatory. We can relax the assumption and allow households to hold risky stocks of the firms not managed by the households. See the discussion at the end of Section 2.1.

5 Households' Asset Distributions in the Steady State

In this model, households' asset distributions in the steady state can be derived analytically. We show below that the asset distributions of entrepreneurs, innate workers, and former entrepreneurs are all Pareto distributions. We also discuss that the asset, income, and consumption distributions of all households follow a Pareto distribution at the upper tail, whose Pareto exponent coincides with that of the asset distribution of entrepreneurs.

5.1 Asset distribution of entrepreneurs

An individual entrepreneur's asset, $\tilde{a}_{e,t}$, if he does not die, evolves as

$$d\ln\tilde{a}_{e,t} = \left(\mu_{ae} - g - \frac{\sigma_{ae}^2}{2}\right)dt + \sigma_{ae}dB_{i,t},$$

where μ_{ae} and σ_{ae} are the drift and diffusion parts of the entrepreneur's asset process, respectively. Because they are constants in the steady state, we omit the time subscript.

The initial asset of entrepreneurs with age t' at period t is $h_{t-t'}$. The relative asset of the entrepreneurs who are alive at t, relative to their initial asset is in a logarithmic expression, $\ln(a_{e,t}/h_{t-t'}) = \ln\tilde{a}_{e,t} - (\ln\tilde{h}_{t-t'} - gt')$ that follows a normal distribution with mean $(\mu_{ae} - \sigma_{ae}^2/2)t'$ and variance $\sigma_{ae}^2 t'$.

We obtain the asset distribution of entrepreneurs by combining the above property with the assumption of constant probability of death. The probability density function of log assets becomes a double-exponential distribution (see Appendix C for the derivations in this section).[6]

$$f_e(\ln\tilde{a}_i) = \begin{cases} f_{e1}(\ln\tilde{a}_i) \equiv \frac{(\nu+p_f)N}{L}\frac{1}{\theta} & \exp\left[-\psi_1(\ln\tilde{a}_i - \ln\tilde{h})\right] \\ & \text{if } \tilde{a}_i \geq \tilde{h}, \\ f_{e2}(\ln\tilde{a}_i) \equiv \frac{(\nu+p_f)N}{L}\frac{1}{\theta} & \exp\left[\psi_2(\ln\tilde{a}_i - \ln\tilde{h})\right] \\ & \text{otherwise}, \end{cases}$$

[6] We normalize the probability density functions of entrepreneurs, innate workers, and former entrepreneurs, $f_e(\ln\tilde{a}_i)$, $f_w(\ln\tilde{a}_i)$, and $f_f(\ln\tilde{a}_i)$, respectively, such that

$$\int_{-\infty}^{\infty}\left\{f_e(\ln\tilde{a}_i) + f_w(\ln\tilde{a}_i) + f_f(\ln\tilde{a}_i)\right\}d(\ln\tilde{a}_i) = 1.$$

where

$$
\begin{aligned}
\psi_1 \equiv & \frac{\mu_{ae} - g - \sigma_{ae}^2/2}{\sigma_{ae}^2} \left(\frac{\theta}{\mu_{ae} - g - \sigma_{ae}^2/2} - 1 \right), \\
\psi_2 \equiv & \frac{\mu_{ae} - g - \sigma_{ae}^2/2}{\sigma_{ae}^2} \left(\frac{\theta}{\mu_{ae} - g - \sigma_{ae}^2/2} + 1 \right), \\
\theta \equiv & \sqrt{2(\nu + p_f)\sigma_{ae}^2 + (\mu_{ae} - g - \sigma_{ae}^2/2)^2}.
\end{aligned}
$$

This result shows that the asset distribution of entrepreneurs follows a double-Pareto distribution (Benhabib et al., 2012 and Toda, 2012), whose Pareto exponent at the upper tail is ψ_1.

5.2 Asset distribution of innate workers

An individual worker's asset, $\tilde{a}_{\ell,t}$, if he does not die, evolves as

$$
d \ln \tilde{a}_{\ell,t} = (\mu_{a\ell} - g)\, dt,
$$

where $\mu_{a\ell}$ is the drift part of the worker's asset process.

Under the asset process, the asset distribution of innate workers becomes

$$
f_w(\ln \tilde{a}_i) = \begin{cases} \frac{\nu L - (\nu + p_f) N}{L} \frac{1}{|\mu_{a\ell} - g|} \exp\left(-\frac{\nu}{\mu_{a\ell} - g} (\ln \tilde{a}_i - \ln \tilde{h}) \right) & \text{if } \frac{\ln \tilde{a}_i - \ln \tilde{h}}{\mu_{a\ell} - g} \geq 0, \\ 0 & \text{otherwise.} \end{cases}
$$

The result shows that the log assets of innate workers follow an exponential distribution, which implies that their assets follow a Pareto distribution. With the parameter values in numerical analysis, the trend growth of workers' assets is close to the trend growth of the economy, that is, $\mu_{a\ell} \approx g$. Then, the detrended assets of the innate workers are concentrated on the level around $\tilde{h}$.

5.3 Asset distribution of former entrepreneurs

The asset distribution of former entrepreneurs depends on the asset distribution of entrepreneurs, the Poisson rate p_f with which each entrepreneur leaves the firm, and the asset process after the entrepreneur becomes a worker.

We can analytically derive the steady state asset distribution of former entrepreneurs. Here, for brevity, we only report the case where $\mu_{a\ell} \geq g$ (for the $\mu_{a\ell} < g$ case, see Appendix C).

$$f_f(\ln \tilde{a}_i) = \begin{cases} \frac{p_f}{\nu - \psi_1(\mu_{a\ell} - g)} f_{e1}(\ln \tilde{a}_i) - \left(\frac{1}{\nu - \psi_1(\mu_{a\ell} - g)} - \frac{1}{\nu + \psi_2(\mu_{a\ell} - g)} \right) p_f f_{e1}(\ln \tilde{h}) \\ \quad \times \exp\left(-\frac{\nu}{\mu_{a\ell} - g} (\ln \tilde{a}_i - \ln \tilde{h}) \right) \quad \text{if } \ln \tilde{a}_i \geq \ln \tilde{h}, \\ \frac{p_f}{\nu + \psi_2(\mu_{a\ell} - g)} f_{e2}(\ln \tilde{a}_i) \quad \text{otherwise.} \end{cases}$$

The probability density function for $\tilde{a}_i \geq \tilde{h}$ consists of two exponential terms. As the asset level increases, the second term, representing the innate workers' distribution, declines faster than the first term, the distribution of entrepreneurs. Therefore, the Pareto exponent of the former entrepreneurs' asset distribution becomes the same as that for entrepreneurs in the tail (the same result applies to the case where $\mu_{a\ell} < g$).

5.4 Pareto exponents of asset and income distributions for all of the households

We make two remarks on the households' asset and income distributions. First, the Pareto exponent at the upper tail of the households' asset distribution is the same as that of the entrepreneurs' asset distribution, ψ_1. This is because, as noted above, the distribution of the smallest Pareto exponent dominates at the upper tail (see e.g., Gabaix, 2009).

Second, in this model, the consumption and income distributions at the upper tail are also Pareto distributions with the same Pareto exponent as that of assets, ψ_1. This is because the consumption and income of a household are proportional to the household's asset level.

6 Numerical Analysis

In this section, we numerically analyze how a reduction in the top marginal tax rate accounts for the evolution of top incomes in recent decades. For this, we assume that an unexpected and permanent tax cut occurs in 1975.

There are three reasons for choosing 1975 as the year of the structural change. First, several empirical studies suggest that inequality has begun to grow since the 1970s (see for example, Katz and Murphy, 1992 and Piketty and Saez, 2003). Second, some political scientists argue that U.S. politics transformed during the 1970s in favor of industries (Hacker and Pierson, 2010), which might have affected entrepreneurs' future expectations on tax rates. Third, the top marginal earned income tax declined from 77% to 50% during the 1970s alone (see Figure 1). This would make CEOs anticipate a subsequent cut in the top ordinary earned income tax, the most important variable in our analysis to account for the evolution of top incomes. They suggest that a structural change occurred during the 1970s.

In our model, a tax cut affects top incomes by changing entrepreneurs' incentive to invest in risky stocks. In the tax parameters calibrated below, after 1975, the tax rate on risky stock τ^e becomes relatively lower than the tax rate on the risk-free asset τ^f. This induces entrepreneurs to increase the share of risky stocks in their asset portfolios. This is the reason why the Pareto exponent declines and the top income share increases in our model.

6.1 Tax rates

We assume that risky stock in our model is a representation of incentive pay, such as employee stock options. Thus, we set the tax on risky stocks τ^e to be equal to the top marginal ordinary income tax that is imposed on top CEOs' pay. On the other hand, the tax on risk-free assets, we assume, is the sum of taxes that investors bear when they hold equities. We calculate the tax on risk-free assets τ^f, by using the equation $1 - \tau^f = (1 - \tau^{\text{cap}})(1 - \tau^{\text{corp}})$, where τ^{cap} and τ^{corp} are the marginal tax rates for capital gains and corporate income, respectively.[7] These tax rates are calibrated by using the top statutory marginal federal tax rates reported in Saez et al. (2012) (see Figure 1 and Table 1).

Insert Figure 1 here.

Insert Table 1 here.

[7] Although we use the capital gains tax primarily because of the availability of data, it can be justified by the following reasoning. We assume that a firm uses all the profit for purchasing shares. Then, the firm pays households money equal to the profit after corporate income tax. The money the households obtain is capital gains, on which capital gains tax is imposed. (Finally, after a part of the after-tax money is paid to financial intermediaries as transaction costs, the households obtain the residual.)

6.2 Calibration

The parameters are chosen to roughly match the annual data. The first five parameters in Table 2 are standard values. For example, we assume for ν that the average length of life after a household begins working is 50 years.

ρ is set to 0.7, implying that 30% of a firm's sales is rent. The value of ρ is lower than the standard value, owing to two reasons. First, our model's treatment of entrepreneurial income is different from the data—in our model, an entrepreneur's income comes mainly from the firm's dividend, whereas in the data, the CEO's pay, in most situations, is categorized in labor income. A lower ρ is chosen to take this into account. Second, if ρ is high, in the situation that entrepreneurs choose $s_{i,t}$ according to (3), the total value of an entrepreneur's risky stocks exceeds the total value of financial assets in the economy. To avoid this, a low ρ should be chosen.

For p_f, we assume that the CEO's average term of office is 20 years. $\ell_{\min}$ is set to unity, that is, the minimum employment level is one person. We assume that $L = 1.0$ and $N = 0.05$. This implies that the average employment of a firm is 20 persons that is consistent with the data reported in Davis et al. (2007). Under the settings, the Pareto exponent of the firm size distribution in the model is $1/(1-0.05) \approx 1.0526$ that is consistent with Zipf's law. Note that under these parameters, for small-sized firms, the value of an entrepreneur's risky stock calculated by (3) exceeds the value of his firm. To resolve this problem, we assume that such an entrepreneur jointly runs a business with other entrepreneurs, such that the asset value of the entrepreneurs' risky stocks does not exceed the value of the joint firms. We assume that the productivity shocks of the joint firms move in the same direction. A possible reason for this assumption is that the productivity shocks are caused by managerial decisions.

For the calibration of firm-level volatility, we consider two cases. In Case A, we use the average firm-level volatility of publicly traded firms. In Case B, we use the average firm-level volatility of both publicly traded and privately held firms. These values are taken from Davis et al. (2007). In each case, the transaction cost of financial intermediaries, ι, is calibrated to match the Pareto exponent in the pre-1975 steady state with 2.4 that is close to the data around 1975. To investigate the extent to which the calibrated ι is reasonable, we compute the model's predictions on the size of the financial sector over GDP, $\iota\left(1 - \frac{\tilde{A}_{e,t}x_t}{\tilde{Q}_t}\right)\tilde{D}_t/\tilde{Y}_t$, under the calibrated ι in Table 3. We find that the model's predictions under the calibrated ι are roughly comparable with the data.

Insert Table 2 here.

Insert Table 3 here.

6.3 Computation of transition dynamics

We compute the Pareto exponent of the household's income (or asset) distribution and the top 1% income share before and after 1975. We assume that before 1975, the economy is in the pre-1975 steady state. In our experiment, taxes change unexpectedly and permanently in 1975, and the economy moves toward the post-1975 steady state.

We model the transition dynamics after 1975 as follows. First, the dynamics of aggregate variables are computed separately. To compute the dynamics of a set of the aggregate variables $\tilde{\mathbf{S}}_t \equiv \mathbf{S}_t/e^{gt} = (\tilde{A}_{e,t}, \tilde{A}_{w,t}, \tilde{A}_{f,t}, \tilde{H}_t, \tilde{K}_t)$ explained in Section 4.1, we need to pin down their initial values. We suppose that when the tax change occurs in 1975, the aggregate capital stock is the same as that in the pre-1975 steady state. For the ease of computation, we also suppose the perfect risk-sharing for the unexpected but verifiable change in the asset values that is caused by the tax change. Then, asset shares of entrepreneurs, innate workers, and former entrepreneurs, $A_{e,1975}/A_{1975}$, $A_{w,1975}/A_{1975}$, $A_{f,1975}/A_{1975}$, respectively, are the same as those in the pre-1975 steady state. The remaining initial variables, $\tilde{A}_{1975}$ and $\tilde{H}_{1975}$ are determined by using the shooting algorithm and the following steps:

1. Set $\tilde{A}_{1975}$. Set also the upper and lower bound of $\tilde{A}_t$, $\tilde{A}_H$ and $\tilde{A}_L$.
   (a) Set $\tilde{H}_{1975}$ and compute the dynamics of aggregate variables as explained in Section 4.1. Stop the computation if $\tilde{A}_t$ hits the upper or lower bound, $\tilde{A}_H$ or $\tilde{A}_L$.
   (b) Update $\tilde{H}_{1975}$ by solving (25) backward, with the terminal condition
   $$\tilde{H}_T = \frac{(1-\alpha)\rho\tilde{p}\tilde{y}^* + \tilde{tr}^*}{\nu + r^{f*} - g},$$
   where the variables with asterisks are those in the post-1975 steady state and $T = \arg\min_t \sqrt{(\tilde{K}_t - \tilde{K}^*)^2 + (\tilde{C}_t - \tilde{C}^*)^2}$.
   (c) Repeat (a) and (b) until $|\tilde{H}^{\mathrm{new}}_{1975} - \tilde{H}^{\mathrm{old}}_{1975}| < \varepsilon$.

2. Repeat the procedure and find the initial value $\tilde{A}_{1975}$ under which the sequence of $\{\tilde{K}_t, \tilde{C}_t\}_t$ converges to the post-1975 steady state.[8]

Note that since $\tilde{C}_t = v\tilde{A}_t$, the above procedure is similar to the shooting algorithm used in standard growth models. In computing the variables used below, we assume that after time T^*, when the dynamics of K_t and C_t are the closest to the post-1975 steady state, the economy switches to the post-1975 steady state.

Next, from the aggregate variables calculated above, we compute the variables related to the entrepreneur's and worker's asset processes, $\mu_{ae,t}$, $\sigma_{ae,t}$, and $\mu_{a\ell,t}$, respectively. Using these variables, we compute the asset (and thus income) distribution at the upper tail. The transition dynamics of the distribution can be computed by numerically solving the Fokker–Planck equations for the asset distributions of entrepreneurs and workers, $f_e(\ln \tilde{a}_{i,t}, t)$ and $f_\ell(\ln \tilde{a}_{i,t}, t) \equiv f_w(\ln \tilde{a}_{i,t}, t) + f_f(\ln \tilde{a}_{i,t}, t)$, respectively, as follows:[9]

$$\begin{aligned}\frac{\partial f_e(\ln \tilde{a}_{i,t}, t)}{\partial t} =& -\left(\mu_{ae,t} - \frac{\sigma_{ae,t}^2}{2} - g\right)\frac{\partial f_e(\ln \tilde{a}_{i,t}, t)}{\partial \ln \tilde{a}_{i,t}} \\ &+ \frac{\sigma_{ae,t}^2}{2}\frac{\partial^2 f_e(\ln \tilde{a}_{i,t}, t)}{\partial (\ln \tilde{a}_{i,t})^2} - (\nu + p_f) f_e(\ln \tilde{a}, t), \\ \frac{\partial f_\ell(\ln \tilde{a}_{i,t}, t)}{\partial t} =& -\left(\mu_{a\ell,t} - g\right)\frac{\partial f_\ell(\ln \tilde{a}_{i,t}, t)}{\partial \ln \tilde{a}_{i,t}} - (\nu - p_f) f_e(\ln \tilde{a}, t).\end{aligned}$$

We impose the boundary conditions that $\lim_{\tilde{a}_{i,t}\to\infty} f_i(\ln \tilde{a}_{i,t}, t) = 0$ and that at the lower bound of $\tilde{a}_{i,t}$, $\tilde{a}_{\mathrm{LB}}$, $f_i(\ln \tilde{a}_{\mathrm{LB}}, t)$ moves linearly during the 50 years from the pre-1975 to the post-1975 steady state.[10]

6.4 Pareto exponent and the top 1% income share

Figures 2 and 3 plot the model's predictions of the Pareto exponent and the top 1% share of income distribution for Case A, together with the data. Data are taken from Alvaredo et al. (2013). For the model's predictions, we plot the two steady states for the pre- and post-1975 periods and the transition path between them.[11]

[8] More specifically, we choose the sequence of $\{\tilde{K}_t, \tilde{C}_t\}_t$ whose distance is the closest to the post-1975 steady state values, $(\tilde{K}^*, \tilde{C}^*)$.

[9] We use the partial differential equations solver in Matlab. We set the 2000 mesh points to $\ln \tilde{a}_{i,t}$ between $\ln \tilde{a}_{\mathrm{LB}}$ and 100 and 500 mesh points to time t between 1975 and 2030.

[10] $\tilde{a}_{\mathrm{LB}}$ is set to be higher than $\tilde{h}$ at the pre- and post-1975 steady states.

[11] The Pareto exponent during the transition path is calculated from the slope of the countercumulative distribution of asset between top 0.1% and top 1%.

We find that the model traces data for the Pareto exponent well. Although ι is set to match the level of the Pareto exponent at the initial steady state, it is non-trivial that the model matches both the level and changes in the Pareto exponent afterward. For example, suppose that we need to set a low (high) ι to match the Pareto exponent at the initial steady state. Then, the changes in the Pareto exponent during the transition become slower (faster) than the data because the volatility of each entrepreneur's asset decreases (increases).

The model also captures the trend in the top 1% share of income after 1975, although the model's prediction is somewhat lower in level than what the data reveal. It is possible that other factors, such as the differences in talents, account for the gap between them.

The corresponding results for Case B are graphed in Figures 4 and 5. The model's transitions of the Pareto exponent and the top 1% share of income become slower than those in Case A. This is because the firm's volatility becomes higher in Case B. This makes $x_{e,t}$ lower by (3), which results in lower volatility of the entrepreneur's asset. This perhaps implies that the lower firm-level volatility in the top firms, where the richest CEOs are employed, is an important factor in understanding the evolution of top incomes.

To take a closer look at the evolution of inequality in the model, in Figure 6, we plot the countercumulative distributions of the household's detrended asset, $\Pr(\tilde{a}_{i,j} > \tilde{a})$, at the pre- and post-1975 steady states and at the transition paths. We find that from a lower asset level, the asset distribution converges to the new stationary distribution at the post-1975 steady state. In other words, the convergence is slower at the wealthiest level. We also find that the convergence is faster in Case A than in Case B that is consistent with the above results.

Insert Figure 2 here.

Insert Figure 3 here.

Insert Figure 4 here.

Insert Figure 5 here.

Insert Figure 6 here.

6.5 Implications of the model

6.5.1 Incentive pay for CEOs

In the real world, CEOs obtain incentive pay, such as stock options, whose value moves along with the performance of the firm. In our model, this is represented by entrepreneurs holding risky stocks of their firms. Here, we discuss whether our formulation is realistic.

Our formulation of CEO pay has a close similarity with those of Edmans et al. (2009) and Edmans et al. (2012). These papers theoretically derive that under the optimal incentive scheme of a CEO in a moral hazard problem, a fraction of the CEO's total assets, denoted by $x_{e,t}$ in our model, is invested in his firm's stocks. Although our model does not take into account the moral hazard problem of CEOs, our model has a similar feature. Edmans et al. (2009) also find evidence that an empirical counterpart of $x_{e,t}$, "percent–percent" incentives, which is a variant of (27) below, is cross-sectionally independent of the firm size.[12] This property is satisfied both in their and our models.

There are also differences between our model and those of Edmans et al. (2009) and Edmans et al. (2012). In their models, only the disutility of effort, a deep parameter, affects the fraction of the entrepreneur's assets invested in his firm's stocks. In our model, however, several factors affect this fraction; for example, an increase in the volatility of the firm value decreases the fraction of the entrepreneur's total assets invested in risky stocks $x_{e,t}$ (see (3)). This prediction is consistent with the evidence surveyed in Frydman and Jenter (2010, Section 2.3).

In our model, changes in taxes also affect the fraction of the entrepreneur's assets invested in his firm's stocks. This is a crucial factor in interpreting the recent evolution of top incomes. After the tax change, top incomes evolve in our model, because it becomes more profitable for CEOs to hold risky stocks. Thus, the tax change induces entrepreneurs' holdings of risky stocks; that is, it induces an increase in $x_{e,t}$ in the post-1975 periods. In the real world, this shows up as the increase in employee stock options. To check the plausibility of our formulation, we compare the model's prediction with the data on incentive pay for CEOs.

[12] The difference between the "percent–percent" incentives and (27) below is that in the "percent–percent" incentives the numerator is "x% increase in the CEO's pay." They are equivalent in the model due to (4) if the CEO's pay is defined by the entrepreneur's consumption as in Edmans et al. (2012).

An empirical counterpart of $x_{e,t}$ is

$$\frac{x\% \text{ increase in the CEO's wealth}}{1\% \text{ increase in firm rate of return}}, \tag{27}$$

because in our model, from (1), it is equal to

$$\frac{d(a_{e,t})/a_{e,t}}{\mu_{q,t}dt + \sigma_{q,t}dB_{e,t}} = x_{e,t}.$$

Unfortunately, a long-term estimate of (27) that covers the pre- and post-1975 periods is not available.

Alternatively, a long-term estimate of a wealth–performance sensitivity measure (referred to as B^I in Edmans et al., 2009),

$$\frac{x\% \text{ increase in the CEO's wealth}}{1\% \text{ increase in firm rate of return}} \times \frac{\text{the CEO's wealth}}{\text{the CEO's pay}}, \tag{28}$$

which is a modification of (27), can be calculated from data in Frydman and Saks (2010).[13] We plot the wealth-performance measure constructed from Frydman and Saks (2010) and the model's counterpart in Figure 7.[14] We find that the wealth–performance measure has increased in the post-1975 period. Our model is qualitatively consistent with the data. The model interprets that this is brought about by the increase in $x_{e,t}$. Quantitatively, in Case A, the model's prediction accounts for the magnitude of the change in the wealth–performance measure occurred in the post-1975 period, although the model does not account for the level of the measure. The opposite results apply for Case B. Of course, our model is not intended to explain the fluctuations in the wealth–performance measure itself, and it cannot explain why these incentives increase around the late 1950s. Further research is needed to understand these empirical facts.

Insert Figure 7 here.

[13]This measure is calculated by dividing the "dollar change in wealth for a 1% increase in firm rate of return" by "total compensation," both of which are taken from Figures 5 and 6 of Frydman and Saks (2010).

[14]The model's counterpart of the wealth-performance measure in (28) is calculated from

$$\frac{d(a_{e,t})/a_{e,t}}{\mu_{q,t}dt + \sigma_{q,t}dB_{e,t}} \frac{a_{e,t}}{\mu_{a,t}a_{e,t} + c_{e,t}} = \frac{x_{e,t}}{\mu_{a,t} + \beta + \nu}.$$

6.5.2 Effect of the tax change on capital accumulation

An important implication of the model is that the tax change does not significantly affect the capital accumulation or capital–output ratio of the economy. This result comes from the property that investment on capital is financed by retained earnings (for details, see Sinn, 1991 and McGrattan and Prescott, 2005). Then, the tax change does not affect the return on stocks $((1-\tau)d_{i,t}dt + dq_{i,t})/q_{i,t}$, because $q_{i,t}$ in the denominator of the equation changes to exactly offset the effect of tax change $(1-\tau)$ in the numerator.

This prediction of the model is in stark contrast to that obtained in previous models of income distribution. However, it is consistent with the facts in the U.S. that the capital–output ratio has not changed significantly over the post-World War II years nor the level of per capita output has increased recently.

6.5.3 Welfare analysis

How has the tax change affected the welfare of households? To determine this, we calculate the utility level of an entrepreneur and an innate worker (that is, a worker from the beginning of his life) in the pre- and post-1975 steady states. Table 4 shows the detrended initial utility level, defined by $\widetilde{V}^i(\tilde{h}, \mathbf{S}) \equiv V^i(h, \mathbf{S}) - gt$ under parameterization of Cases A and B. (for details of the derivations, see Appendix D).

Not surprisingly, the utility level of an innate worker becomes lower in the post-1975 steady state under Cases A and B parameterizations, whereas that of an entrepreneur becomes higher under Case A parameterization. These results are consistent with the view that the rich have benefited from the tax change at the expense of the poor. Interestingly, under Case B parameterization, the utility level of an entrepreneur also becomes lower in the post-1975 steady state. The result seems to stem from the property that taxes and transfers in the model play the role of an insurance device. Under Case B parameterization, where firm-level volatility is high, the disappearance of the insurance device has a detrimental effect on not only workers, but also entrepreneurs.

Insert Table 4 here.

7 Conclusion

We have proposed a model of asset and income inequalities that explains both Zipf's law of firms and Pareto's law of incomes from the idiosyncratic productivity shocks of firms. Empirical studies show that the Pareto exponent of income varies over time, whereas Zipf's law of firm size is quite stable. This paper consistently explains these distributions with an analytically tractable model. We derive closed-form expressions for the stationary distributions of firm size and individual income. The transition dynamics of those distributions are also explicitly derived and are then used for numerical analysis.

Our model features an entrepreneur who can invest in his own firm as well as in risk-free assets. The entrepreneur incurs a substantial transaction cost if he diversifies the risk of his portfolio returns. When a tax on risky returns is reduced, the entrepreneur increases the share of his own firm. This, in turn, increases the variance of his portfolio returns, resulting in a wider dispersion of wealth among entrepreneurs.

By calibrating the model, we have analyzed to what extent the changes in tax rates account for the recent evolution of top incomes in the U.S. We find that the model matches the decline in the Pareto exponent of income distribution and the trend in the top 1% income share. There remain some discrepancies between the model and data. For example, the model's prediction of the top 1% share is somewhat lower than the data. Further research is needed to understand the causes of such discrepancies.

References

Alvaredo, Facundo, Anthony B. Atkinson, Thomas Piketty, and Emmanuel Saez (2013) "The World Top Incomes Database." Data are extracted at 25/11/2013.

Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez (2011) "Top Incomes in the Long Run of History," *Journal of Economic Literature*, Vol. 49, No. 1, pp. 3–71.

Bakija, Jon, Adam Cole, and Bradley T. Heim (2012) "Jobs and Income Growth of Top Earners and the Causes of Changing Income Inequality: Evidence from U.S. Tax Return Data," mimeo.

Benhabib, Jess, Alberto Bisin, and Shenghao Zhu (2011) "The Distribution of Wealth and Fiscal Policy in Economies with Finitely Lived Agents," *Econometrica*, Vol. 79, No. 1, pp. 123–157.

——— (2012) "The Distribution of Wealth in the Blanchard-Yaari Model," mimeo.

Blanchard, Olivier J (1985) "Debt, Deficits, and Finite Horizons," *Journal of Political Economy*, Vol. 93, No. 2, pp. 223–47.

Campbell, John Y. and Luis M. Viceira (2002) *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*: Oxford University Press.

Davis, Steven J., John Haltiwanger, Ron Jarmin, and Javier Miranda (2007) "Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms," in *NBER Macroeconomics Annual 2006, Volume 21*: National Bureau of Economic Research, Inc, pp. 107–180.

Edmans, Alex, Xavier Gabaix, and Augustin Landier (2009) "A Multiplicative Model of Optimal CEO Incentives in Market Equilibrium," *Review of Financial Studies*, Vol. 22, No. 12, pp. 4881–4917.

Edmans, Alex, Xavier Gabaix, Tomasz Sadzik, and Yuliy Sannikov (2012) "Dynamic CEO Compensation," *Journal of Finance*, Vol. 67, No. 5, pp. 1603–1647.

Frydman, Carola and Dirk Jenter (2010) "CEO Compensation," *Annual Review of Financial Economics*, Vol. 2, No. 1, pp. 75–102.

Frydman, Carola and Raven E. Saks (2010) "Executive Compensation: A New View from a Long-Term Perspective, 1936–2005," *Review of Financial Studies*, Vol. 23, No. 5, pp. 2099–2138.

Gabaix, Xavier (2009) "Power Laws in Economics and Finance," *Annual Review of Economics*, Vol. 1, No. 1, pp. 255–294.

Gabaix, Xavier and Augustin Landier (2008) "Why Has CEO Pay Increased So Much?" *Quarterly Journal of Economics*, Vol. 123, No. 1, pp. 49–100.

Greenwood, Robin and David Scharfstein (2013) "The Growth of Finance," *Journal of Economic Perspectives*, Vol. 27, No. 2, pp. 3–28.

Hacker, Jacob S. and Paul Pierson (2010) *Winner-Take-All Politics: How Washington Made the Rich Richer–and Turned Its Back on the Middle Class*: Simon and Schuster.

Jones, Charles I. and Jihee Kim (2012) "Exploring the Dynamics of Top Income Inequality," mimeo.

——— (2013) “A Schumpeterian Model of Top Income Inequality,” mimeo.

Kaplan, Steven N. and Joshua Rauh (2013) “It's the Market: The Broad-Based Rise in the Return to Top Talent,” *Journal of Economic Perspectives*, Vol. 27, No. 3, pp. 35–56.

Katz, Lawrence F and Kevin M Murphy (1992) “Changes in Relative Wages, 1963-1987: Supply and Demand Factors,” *Quarterly Journal of Economics*, Vol. 107, No. 1, pp. 35–78.

Kim, Jihee (2013) “The Effect of the Top Marginal Tax Rate on Top Income Inequality,” mimeo.

Luttmer, Erzo G. J. (2007) “Selection, Growth, and the Size Distribution of Firms,” *Quarterly Journal of Economics*, Vol. 122, No. 3, pp. 1103–1144, 08.

Luttmer, Erzo G.J. (2010) “Models of Growth and Firm Heterogeneity,” *Annual Review of Economics*, Vol. 2, No. 1, pp. 547–576.

McGrattan, Ellen R. and Edward C. Prescott (2005) “Taxes, Regulations, and the Value of U.S. and U.K. Corporations,” *Review of Economic Studies*, Vol. 72, No. 3, pp. 767–796.

Merton, Robert C. (1969) “Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case,” *Review of Economics and Statistics*, Vol. 51, No. 3, pp. 247–57.

——— (1971) “Optimum Consumption and Portfolio Rules in a Continuous-time Model,” *Journal of Economic Theory*, Vol. 3, No. 4, pp. 373–413.

——— (1973) “An Intertemporal Capital Asset Pricing Model,” *Econometrica*, Vol. 41, No. 5, pp. 867–87.

Mizuno, Takayuki, Atushi Ishikawa, Shouji Fujimoto, and Tsutomu Watanabe (2012) “Power Laws in Firm Productivity,” *Progress of Theoretical Physics Supplement*, Vol. 194, pp. 122-134.

Nirei, Makoto (2009) “Pareto Distributions in Economic Growth Models,” IIR Working Paper 09-05, Institute of Innovation Research, Hitotsubashi University.

Nirei, Makoto and Wataru Souma (2007) “A Two Factor Model Of Income Distribution Dynamics,” *Review of Income and Wealth*, Vol. 53, No. 3, pp. 440–459.

Piketty, Thomas, Emmanuel Saez, and Stefanie Stantcheva (2011) "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities," NBER Working Papers 17616, National Bureau of Economic Research, Inc.

Piketty, Thomas and Emmanuel Saez (2003) "Income Inequality in the United States, 1913-1998," *Quarterly Journal of Economics*, Vol. 118, No. 1, pp. 1–39.

Rossi-Hansberg, Esteban and Mark L. J. Wright (2007) "Establishment Size Dynamics in the Aggregate Economy," *American Economic Review*, Vol. 97, No. 5, pp. 1639–1666.

Saez, Emmanuel, Joel Slemrod, and Seth H. Giertz (2012) "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," *Journal of Economic Literature*, Vol. 50, No. 1, pp. 3–50.

Sinn, Hans-Werner (1991) "Taxation and the Cost of Capital: The 'Old' View, the 'New' View, and Another View," in *Tax Policy and the Economy, Volume 5*: National Bureau of Economic Research, Inc, pp. 25–54.

Toda, Alexis Akira (2012) "Asset Pricing and Wealth Distribution with Heterogeneous Investment Returns," mimeo.

A Derivations for the household's problem

This appendix shows the derivations of the household problem in Section 2.1. As shown in Section 4.1, the aggregate dynamics of the model is described by $\mathbf{S}_t$, whose evolution can be written as

$$d\mathbf{S}_t = \mu_{\mathbf{S}}(\mathbf{S}_t)dt.$$

By Ito's formula, $V^i(a_{i,t}, \mathbf{S}_t)$ is rewritten as follows:

$$\begin{aligned} dV^i(a_{i,t}, \mathbf{S}_t) =& \frac{\partial V_t^i}{\partial a_{i,t}} da_{i,t} + \frac{1}{2}\frac{\partial^2 V_t^i}{\partial a_{i,t}^2}(da_{i,t})^2 + \frac{\partial V_t^i}{\partial \mathbf{S}_t} \cdot d\mathbf{S}_t \\ &+ \left(V^\ell(a_{i,t}, \mathbf{S}_t) - V^i(a_{i,t}, \mathbf{S}_t)\right) dJ_{i,t}, \end{aligned}$$

where $J_{i,t}$ is the Poisson jump process that describes the probability of an entrepreneur leaving his firm and becoming a worker.

$$dJ_{i,t} = \begin{cases} 0 & \text{with probability } 1 - p_f dt \\ 1 & \text{with probability } p_f dt. \end{cases}$$

Thus,

$$\frac{\mathsf{E}_t[dV_t^i]}{dt} = \mu_{a,t} a_{i,t} \frac{\partial V_t^i}{\partial a_{i,t}} + \frac{(\sigma_{a,t} a_{i,t})^2}{2} \frac{\partial^2 V_t^i}{\partial a_{i,t}^2} + \mu_{\mathbf{S}}'(\mathbf{S}_t) \cdot \frac{\partial V_t^i}{\partial \mathbf{S}_t} + p_f \left(V_t^\ell - V_t^i \right).$$

Substituting in (2), we obtain a Hamilton–Jacobi–Bellman equation as follows:

$$\begin{aligned} 0 &= \max_{c_{i,t}, x_{i,t}} \ln c_{i,t} - (\beta + \nu) V_t^i + \mu_{a,t} a_{i,t} \frac{\partial V_t^i}{\partial a_{i,t}} + \frac{(\sigma_{a,t} a_{i,t})^2}{2} \frac{\partial^2 V_t^i}{\partial a_{i,t}^2} \\ &+ \mu_{\mathbf{S}}'(\mathbf{S}_t) \cdot \frac{\partial V_t^i}{\partial \mathbf{S}_t} + p_f \left(V_t^\ell - V_t^i \right) \\ &= \max_{c_{i,t}, x_{i,t}} \ln c_{i,t} - (\beta + \nu) V_t^i + \frac{\sigma_{q,t}^2}{2} x_{i,t}^2 a_{i,t}^2 \frac{\partial^2 V_t^i}{\partial a_{i,t}^2} \\ &+ ((\nu + \mu_{q,t}) x_{i,t} a_{i,t} + (\nu + r_t^f)(1 - x_{i,t}) a_{i,t} - c_{i,t}) \frac{\partial V_t^i}{\partial a_{i,t}}, \\ &+ \mu_{\mathbf{S}}'(\mathbf{S}_t) \cdot \frac{\partial V_t^i}{\partial \mathbf{S}_t} + p_f \left(V_t^\ell - V_t^i \right). \end{aligned} \tag{29}$$

The FOCs with respect to $c_{i,t}$ and $x_{i,t}$ are summarized as follows:

$$c_{i,t}^{-1} = \frac{\partial V_t^i}{\partial a_{i,t}}, \tag{30}$$

$$x_{i,t} = \begin{cases} -\frac{\partial V_t^i / \partial a_{i,t}}{\left(\partial^2 V_t^i / \partial a_{i,t}^2\right) a_{i,t}} \frac{\mu_{q,t} - r_t^f}{\sigma_{q,t}^2}, & \text{if } i = e, \\ 0, & \text{otherwise.} \end{cases} \tag{31}$$

Moreover, (29) has to satisfy the transversality condition (5).

Following Merton (1969) and Merton (1971), this problem is solved by the following value function and linear policy functions:

$$\begin{aligned} V_t^i &= B_t^i \ln a_{i,t} + H^i(\mathbf{S}_t), \\ c_{i,t} &= v_{i,t} a_{i,t}, \\ q_{i,t} s_{i,t} &= x_{i,t} a_{i,t}, \\ b_{i,t} &= (1 - x_{i,t}) a_{i,t} - h_t. \end{aligned} \tag{32}$$

We obtain this solution by guess–and–verify. The FOC (30) becomes

$$(v_{i,t})^{-1} = B_t^i.$$

Condition (31) is rewritten as

$$x_{i,t} = \begin{cases} \frac{\mu_{q,t} - r_t^f}{\sigma_{q,t}^2}, & \text{if } i = e, \\ 0, & \text{otherwise.} \end{cases}$$

Substituting these results into (29), we find that

$$v_{i,t} = \beta + \nu.$$

B Derivations for the firm's problem

B.1 Derivations of FOCs of the firm's problem

This appendix shows the derivations of the firm's problem described in Section 2.2.2. $q_{j,t}$ is a function of $k_{j,t}$, $z_{j,t}$, and the aggregate dynamics $\mathbf{S}_t$ (see Appendix A). By applying Ito's formula to $q_{j,t}$, we obtain

$$\begin{aligned} dq(k_{j,t}, z_{j,t}, \mathbf{S}_t) &= \left(\frac{\partial q_{j,t}}{\partial z_{j,t}} dz_{j,t} + \frac{\partial q_{j,t}}{\partial k_{j,t}} dk_{j,t} + \frac{\partial q_{j,t}}{\partial \mathbf{S}_t} \cdot d\mathbf{S}_t \right) + \frac{1}{2} \frac{\partial^2 q_{j,t}}{\partial z_{j,t}^2} (dz_{j,t})^2 \\ &= \left(\mu_z \frac{\partial q_{j,t}}{\partial z_{j,t}} + \frac{1}{2} \sigma_z^2 \frac{\partial^2 q_{j,t}}{\partial z_{j,t}^2} \right) dt + \frac{\partial q_{j,t}}{\partial k_{j,t}} dk_{j,t} + \mu_{\mathbf{S}}'(\mathbf{S}_t) \cdot \frac{\partial q_{j,t}}{\partial \mathbf{S}_t} \\ &\quad + \sigma_z \frac{\partial q_{j,t}}{\partial z_{j,t}} dB_{j,t}. \end{aligned}$$

The FOCs for $\ell_{j,t}$ and $dk_{j,t}$ are

$$\begin{aligned}(1-\tau^f-\iota) &= \frac{\partial q_{j,t}}{\partial k_{j,t}},\\ w_t &= \frac{\partial p_{j,t}y_{j,t}}{\partial \ell_{j,t}}.\end{aligned}$$

By the envelope theorem,

$$r_t^f \frac{\partial q_{j,t}}{\partial k_{j,t}}dt = (1-\tau^f-\iota)\left(\frac{\partial p_{j,t}y_{j,t}}{\partial k_{j,t}}dt - \delta dt\right).$$

By rearranging the equation, we obtain

$$r_t^f = \frac{\partial p_{j,t}y_{j,t}}{\partial k_{j,t}} - \delta.$$

B.2 Derivations of the firm-side variables

This appendix shows the derivations of the firm-side variables described in Section 3.1. From (9),

$$w_t = (1-\alpha)\rho\left(\frac{Y_t}{N}\right)^{1-\rho} z_{j,t}^{\rho} k_{j,t}^{\alpha\rho} \ell_{j,t}^{(1-\alpha)\rho-1}.$$

Rewriting this,

$$\ell_{j,t} = \left(\frac{(1-\alpha)\rho}{w_t}\left(\frac{Y_t}{N}\right)^{1-\rho} z_{j,t}^{\rho} k_{j,t}^{\alpha\rho}\right)^{\frac{1}{1-(1-\alpha)\rho}}. \tag{33}$$

On the other hand, from (8),

$$\mathsf{MPK}_t = \alpha\rho\left(\frac{Y_t}{N}\right)^{1-\rho} z_{j,t}^{\rho} k_{j,t}^{\alpha\rho-1} \ell_{j,t}^{(1-\alpha)\rho}. \tag{34}$$

By substituting (33) into (34) and rearranging,

$$k_{j,t}^{\frac{\alpha\rho}{1-(1-\alpha)\rho}} = \left(\frac{\alpha\rho}{\mathsf{MPK}_t}\left(\frac{Y_t}{N}\right)^{1-\rho}\right)^{\frac{\alpha\rho}{1-\rho}} \left(\frac{(1-\alpha)\rho}{w_t}\left(\frac{Y_t}{N}\right)^{1-\rho}\right)^{\frac{\alpha\rho(1-\alpha)\rho}{(1-\rho)(1-(1-\alpha)\rho)}} z_{j,t}^{\eta\frac{\alpha\rho}{1-\rho}}, \tag{35}$$

where $\eta \equiv \frac{\rho}{1-(1-\alpha)\rho}$. Substituting (35) into (33),

$$\ell_{j,t} = \left(\frac{\alpha\rho}{\mathsf{MPK}_t} \left(\frac{Y_t}{N}\right)^{1-\rho} \right)^{\frac{\alpha\rho}{1-\rho}} \left(\frac{(1-\alpha)\rho}{w_t} \left(\frac{Y_t}{N}\right)^{1-\rho} \right)^{\frac{1-\alpha\rho}{1-\rho}} z_{j,t}^{\phi} \tag{36}$$

By substituting this equation into the labor market condition (11) and rearranging,

$$\left(\frac{\alpha\rho}{\mathsf{MPK}_t} \left(\frac{Y_t}{N}\right)^{1-\rho} \right)^{\frac{\alpha\rho}{1-\rho}} \left(\frac{(1-\alpha)\rho}{w_t} \left(\frac{Y_t}{N}\right)^{1-\rho} \right)^{\frac{1-\alpha\rho}{1-\rho}} = \frac{L}{N} \frac{1}{\overline{\mathsf{E}}\left\{z_{j,t}^{\phi}\right\}}, \tag{37}$$

$$\text{or,} \quad \left(\frac{(1-\alpha)\rho}{w_t} \left(\frac{Y_t}{N}\right)^{1-\rho} \right)^{\frac{(1-\alpha)\rho}{1-\rho}} = \left\{ \left(\frac{\alpha\rho}{\mathsf{MPK}_t} \left(\frac{Y_t}{N}\right)^{1-\rho} \right)^{\frac{-\alpha\rho}{1-\rho}} \frac{L}{N} \frac{1}{\overline{\mathsf{E}}\left\{z_{j,t}^{\phi}\right\}} \right\}^{\frac{(1-\alpha)\rho}{1-\alpha\rho}}. \tag{38}$$

Here, $\overline{\mathsf{E}}$ is the operator of the cross-sectional average of all firms. Then, substituting (37) into (36),

$$\ell_{j,t} = \frac{L}{N} \left(\frac{z_{j,t}^{\phi}}{\overline{\mathsf{E}}\left\{z_{j,t}^{\phi}\right\}} \right). \tag{39}$$

Rewriting (35),

$$k_{j,t} = \left(\frac{\alpha\rho}{\mathsf{MPK}_t} \left(\frac{Y_t}{N}\right)^{1-\rho} \right)^{\frac{1-(1-\alpha)\rho}{1-\rho}} \left(\frac{(1-\alpha)\rho}{w_t} \left(\frac{Y_t}{N}\right)^{1-\rho} \right)^{\frac{(1-\alpha)\rho}{1-\rho}} z_{j,t}^{\phi}. \tag{40}$$

Substituting (38) into (40),

$$k_{j,t} = \left(\frac{\alpha\rho}{\mathsf{MPK}_t} \left(\frac{Y_t}{N}\right)^{1-\rho} \right)^{\frac{1}{1-\alpha\rho}} \left(\frac{L}{N}\right)^{\frac{(1-\alpha)\rho}{1-\alpha\rho}} \left(\frac{z_{j,t}^{\phi}}{\overline{\mathsf{E}}\left\{z_{j,t}^{\phi}\right\}^{\frac{(1-\alpha)\rho}{1-\alpha\rho}}} \right). \tag{41}$$

Next, we derive Y_t. Substituting (39) and (41) into $y_{j,t} = z_{j,t} k_{j,t}^{\alpha} \ell_{j,t}^{1-\alpha}$ and rearranging,

$$y_{j,t} = \left(\frac{\alpha\rho}{\mathsf{MPK}_t} \left(\frac{Y_t}{N}\right)^{1-\rho} \right)^{\frac{\alpha}{1-\alpha\rho}} \left(\frac{L}{N}\right)^{\frac{1-\alpha}{1-\alpha\rho}} \left(\frac{z_{j,t}^{\frac{1}{1-\rho}}}{\overline{\mathsf{E}}\left\{z_{j,t}^{\phi}\right\}^{\frac{(1-\alpha)}{1-\alpha\rho}}} \right).$$

Substituting this equation into $Y_t = \left(\int_0^N \left(\frac{1}{N}\right)^{1-\rho} y_{j,t}^{\rho} dj\right)^{\frac{1}{\rho}}$,

$$\left(\frac{Y_t}{N}\right)^{1-\rho} = \left(\frac{\alpha\rho}{\mathsf{MPK}_t}\right)^{\frac{\alpha(1-\rho)}{1-\alpha}} \left(\frac{L}{N}\right)^{1-\rho} \overline{\mathsf{E}}\left\{z_{j,t}^{\phi}\right\}^{(1-\rho)\left[\frac{1-\alpha\rho}{(1-\alpha)\rho}-1\right]}. \tag{42}$$

Substituting (42) into (41),

$$\begin{aligned} k_{j,t} &= \left(\frac{\alpha\rho}{\mathsf{MPK}_t}\right)^{\frac{1}{1-\alpha}} \overline{\mathsf{E}}\left\{z_{j,t}^{\phi}\right\}^{\frac{1}{\phi}\frac{1}{1-\alpha}} \left(\frac{L}{N}\right) \left(\frac{z_{j,t}^{\phi}}{\overline{\mathsf{E}}\left\{z_{j,t}^{\phi}\right\}}\right) \\ &= \left(\frac{\alpha\rho}{\mathsf{MPK}_t}\right)^{\frac{1}{1-\alpha}} \overline{\mathsf{E}}\left\{z_{j,t}^{\phi}\right\}^{\frac{1}{\phi}\frac{1}{1-\alpha}} \ell_{j,t}. \end{aligned} \tag{43}$$

Substituting (39) and (43) into (42),

$$p_{j,t}y_{j,t} = Y_t^{1-\rho} y_{j,t}^{\rho} \tag{44}$$

$$\begin{aligned} &= \left(\frac{\alpha\rho}{\mathsf{MPK}_t}\right)^{\frac{\alpha}{1-\alpha}} \overline{\mathsf{E}}\left\{z_{j,t}^{\phi}\right\}^{\frac{1}{\phi}\frac{1}{1-\alpha}} \left(\frac{L}{N}\right) \left(\frac{z_{j,t}^{\phi}}{\overline{\mathsf{E}}\left\{z_{j,t}^{\phi}\right\}}\right) \\ &= \left(\frac{\alpha\rho}{\mathsf{MPK}_t}\right)^{\frac{\alpha}{1-\alpha}} \overline{\mathsf{E}}\left\{z_{j,t}^{\phi}\right\}^{\frac{1}{\phi}\frac{1}{1-\alpha}} \ell_{j,t}. \end{aligned} \tag{45}$$

Rewriting (39),

$$\ell_{j,t} = \overline{\ell}_t\ z_{j,t}^{\phi}, \text{ where } \overline{\ell}_t \equiv \left(\frac{L/N}{\overline{\mathsf{E}}\left\{z_{j,t}^{\phi}\right\}}\right).$$

Rewriting (45),

$$p_{j,t}y_{j,t} = \overline{py}_t\overline{\ell}_t\ z_{j,t}^{\phi}, \text{ where } \overline{py}_t \equiv \left(\frac{\alpha\rho}{\mathsf{MPK}_t}\right)^{\frac{\alpha}{1-\alpha}} \overline{\mathsf{E}}\left\{z_{j,t}^{\phi}\right\}^{\frac{1}{\phi}\frac{1}{1-\alpha}}.$$

Rewriting (43),

$$k_{j,t} = \overline{k}_t\overline{\ell}_t z_{j,t}^{\phi}, \text{ where } \overline{k}_t \equiv \left(\frac{\alpha\rho}{\mathsf{MPK}_t}\overline{\mathsf{E}}\left\{z_{j,t}^{\phi}\right\}^{\frac{1}{\phi}}\right)^{\frac{1}{1-\alpha}}. \tag{46}$$

From (46),

$$\begin{aligned} dk_{j,t} &= d(\overline{k}_t\overline{\ell}_t z_{j,t}^{\phi}) \\ &= \frac{d\overline{k}_t\overline{\ell}_t}{dt} z_{j,t}^{\phi} dt + \overline{k}_t\overline{\ell}_t d\left(z_{j,t}^{\phi}\right). \end{aligned}$$

Note that

$$d\left(z_{j,t}^{\phi}\right) = \left\{\phi\mu_z + \phi\left(\phi-1\right)\frac{\sigma_z^2}{2}\right\} z_{j,t}^{\phi} dt + \phi\sigma_z z_{j,t}^{\phi} dB_{j,t}.$$

Then,

$$\begin{aligned} dk_{j,t} &= d(\overline{k}_t\overline{\ell}_t z_{j,t}^{\phi}) \\ &= \frac{d\overline{k}_t\overline{\ell}_t}{dt} z_{j,t}^{\phi} dt + \overline{k}_t\overline{\ell}_t d\left(z_{j,t}^{\phi}\right) \\ &= k_{j,t}\left\{\mu_{k,t} dt + \phi\sigma_z dB_{j,t}\right\}. \end{aligned}$$

Here,

$$\mu_{k,t} \equiv g - \frac{1}{1-\alpha}\frac{dr_t^f/dt}{\mathsf{MPK}_t} + \phi\left\{(\mu_z - g_z) + (\phi-1)\frac{\sigma_z^2}{2}\right\}.$$

g_z is the growth rate of $\overline{\mathsf{E}}\left\{z_{j,t}^{\phi}\right\}$ and $g = g_z/(1-\alpha)$.

$d_{j,t}dt$ is computed by substituting these results into the following relationship:

$$\begin{aligned} d_{j,t}dt =& (p_{j,t}y_{j,t} - w_t\ell_{j,t} - \delta k_{j,t})dt - dk_{j,t} \\ =& (1-(1-\alpha)\rho)p_{j,t}y_{j,t}dt - \delta k_{j,t}dt - dk_{j,t}. \end{aligned}$$

Then, $d_{j,t}dt$ is rewritten as follows:

$$d_{j,t}dt = \overline{d}_t\overline{\ell}_t z_{j,t}^{\phi} dt - \left\{\phi\sigma_z dB_{j,t}\right\}\overline{k}_t\overline{\ell}_t z_{j,t}^{\phi},$$

$$\text{where } \overline{d}_t \equiv (1-(1-\alpha)\rho)\overline{py}_t - (\delta + \mu_{k,t})\,\overline{k}_t.$$

By aggregating the above equation and detrending by e^{gt}, we obtain

$$\tilde{D}_t = (1-(1-\alpha)\rho)\tilde{Y}_t - (\delta + \mu_{\tilde{K}})\tilde{K}_t - \frac{d\tilde{K}_t}{dt}, \text{ where } \mu_{\tilde{K}} \equiv g + \phi\left\{(\mu_z - g_z) + (\phi - 1)\frac{\sigma_z^2}{2}\right\}.$$

Here, we use the property that

$$\frac{1}{1-\alpha}\frac{dr_t^f/dt}{\mathsf{MPK}_t} = \frac{d\tilde{K}_t}{dt}/\tilde{K}_t.$$

B.3 Returns on risky stocks

This appendix explains the derivation of the returns on risky stocks described in Sections 3.1 and 4.1. Multiplying (6) by $e^{-\int_t^u r_s^f ds}$ and integrating,[15] we obtain

$$q_{j,t} = \mathsf{E}_t\left[\int_t^{\infty}(1-\tau^f-\iota)d_{j,u}e^{-\int_t^u r_s^f ds}du\right].$$

By further rearranging the above equation,

$$q_{j,t} = \int_t^{\infty}(1-\tau^f-\iota)e^{-\int_t^u r_s^f ds}\,\mathsf{E}_t\left[d_{j,u}\right]du.$$

Because

$$\begin{aligned}
\mathsf{E}_t[d_{j,u}] =& \overline{d}_u\overline{\ell}_u\,\mathsf{E}_t[z_{j,u}^{\phi}] \\
=& \overline{d}_t\overline{\ell}_t\frac{\overline{d}_u\overline{\ell}_u}{\overline{d}_t\overline{\ell}_t} \times \exp\left\{\int_t^u\left(\phi\mu_z + \phi\left(\phi-1\right)\frac{\sigma_z^2}{2}\right)ds\right\}\cdot z_{j,t}^{\phi} \\
=& \overline{d}_t\overline{\ell}_t z_{j,t}^{\phi}\exp\left\{\int_t^u\left(\frac{d\ln(\overline{d}_s\overline{\ell}_s)}{ds} + \phi\mu_z + \phi\left(\phi-1\right)\frac{\sigma_z^2}{2}\right)ds\right\} \\
=& \overline{d}_t\overline{\ell}_t z_{j,t}^{\phi}\exp\left\{\int_t^u \mu_{d,s}ds\right\}, \text{ where } \mu_{d,t} \equiv \frac{d\ln(\overline{d}_t\overline{\ell}_t)}{dt} + \phi\left(\mu_z + (\phi-1)\frac{\sigma_z^2}{2}\right).
\end{aligned}$$

[15] The Ito process version of integration by parts

$$\int_t^T X_{j,s}dY_{j,s} = X_{j,T}Y_{j,T} - X_{j,t}Y_{j,t} - \int_t^T Y_{j,s}dX_{j,s} - \int_t^T dX_{j,s}dY_{j,s}$$

is used here. Define $\Delta_{t,u} \equiv e^{-\int_t^u r_s^f ds}$. Then,

$$\int_t^{\infty}\Delta_{t,u}dq_{j,u} = q_{j,u}\Delta_{t,u}\mid_t^{\infty} - \int_t^{\infty}q_{j,u}(-r_u^f)\Delta_{t,u}du$$

Therefore,

$$q_{j,t} = \overline{q}_t \overline{\ell}_t z_{j,t}^{\phi}, \quad \text{where } \overline{q}_t \equiv (1-\tau^f-\iota)\overline{d}_t \int_t^\infty \exp\left\{-\int_t^u (r_s^f - \mu_{d,s})ds\right\} du.$$

Then,

$$\begin{aligned} dq_{j,t} =& q_{j,t}\frac{d\ln(\overline{d}_t\overline{\ell}_t)}{dt}dt + q_{j,t}\frac{d(z_{j,t}^{\phi})}{z_{j,t}^{\phi}} + q_{j,t}\frac{-1+(r_t^f-\mu_{d,t})\int_t^\infty \exp\left\{-\int_t^u (r_s^f-\mu_{d,s})ds\right\}du}{\int_t^\infty \exp\left\{-\int_t^u (r_s^f-\mu_{d,s})ds\right\}du}dt \\ =& \left\{-(1-\tau^f-\iota)\overline{d}_t\overline{\ell}_t z_{j,t}^{\phi} + r_t^f q_{j,t}\right\}dt + q_{j,t}\phi\sigma_z dB_{j,t}. \end{aligned}$$

By using $d_{j,t} = \overline{d}_t\overline{\ell}_t z_{j,t}^{\phi}$, the return of a risky stock is

$$\begin{aligned} \frac{(1-\tau^e)d_{j,t}+dq_{j,t}}{q_{j,t}} =& \left\{\left(\frac{1-\tau^e}{(1-\tau^f-\iota)}-1\right)\frac{1}{\int_t^\infty \exp\left\{-\int_t^u (r_s^f-\mu_{d,s})ds\right\}du} + r_t^f\right\}dt \\ &+ \phi\sigma_z\left\{1-\left(\frac{1-\tau^e}{(1-\tau^f-\iota)}\right)\frac{\overline{k}_t}{\overline{d}_t}\frac{1}{\int_t^\infty \exp\left\{-\int_t^u (r_s^f-\mu_{d,s})ds\right\}du}\right\}dB_{j,t}. \end{aligned}$$

Note that if $(r_t^f - \mu_{d,t})$ is constant as in the steady state, $\int_t^\infty \exp\left\{-\int_t^u (r_s^f-\mu_{d,s})ds\right\}du = 1/(r^f-\mu_d)$ and

$$q_{j,t} = \frac{(1-\tau^f-\iota)\overline{d}_t\overline{\ell}_t z_{j,t}^{\phi}}{r^f-\mu_d}.$$

We need to know the value of $\int_t^\infty \exp\left\{-\int_t^u (r_s^f-\mu_{d,s})ds\right\}du$ to compute the return on risky stocks. We calculate the value as follows. Integrating (15), we obtain

$$\int_t^\infty \exp\left\{-\int_t^u (r_s^f-\mu_{d,s})ds\right\}du = \frac{Q_t}{(1-\tau^f-\iota)\overline{d}_t L}.$$

If we know the value of $Q_t = A_t - H_t$ and $(1-\tau^f-\iota)\overline{d}_t L$, we can calculate the value of $\int_t^\infty \exp\left\{-\int_t^u (r_s^f-\mu_{d,s})ds\right\}du$.

B.4 Derivations of the restructuring

This appendix shows the derivations of the restructuring described in Section 3.2. Let $\tilde{z}_{j,t}$ be the firm's productivity level after selling a part of the firm's assets to restructuring firms, detrended by $e^{g_z t}$. Then,

$Q_{\text{restructuring},t+dt}$ is written as follows:

$$Q_{\text{restructuring},t+dt} = N\overline{q}_{t+dt}\overline{\ell}_{t+dt}e^{g_z t}\overline{\mathsf{E}}\left\{\tilde{z}^{\phi}_{\min} - \tilde{z}^{\phi}_{j,t+dt}\middle|\tilde{z}_{j,t+dt} \leq \tilde{z}_{\min}\right\}.$$

Here, $\overline{\mathsf{E}}\left\{\tilde{z}^{\phi}_{\min} - \tilde{z}^{\phi}_{j,t+dt}\middle|\tilde{z}_{j,t+dt} \leq \tilde{z}_{\min}\right\}$ is the expectation of $\tilde{z}^{\phi}_{\min} - \tilde{z}^{\phi}_{j,t+dt}$ conditional on $\tilde{z}_{j,t+dt}$ being lower than $\tilde{z}_{\min}$. Because the evolution of $\tilde{z}_{j,t}$ follows (18) and the distribution follows (19),

$$\begin{aligned}
\overline{\mathsf{E}}\left\{\tilde{z}^{\phi}_{\min} - \tilde{z}^{\phi}_{j,t+dt}\middle|\tilde{z}_{j,t+dt} \leq \tilde{z}_{\min}\right\} = &\int_{\ln \tilde{z}_{\min}}^{\infty} d(\ln \tilde{z}_{j,t}) \int_{-\infty}^{\ln \tilde{z}_{\min}} d(\ln \tilde{z}_{j,t+dt}) \\
&\left(\tilde{z}^{\phi}_{\min} - \tilde{z}^{\phi}_{j,t+dt}\right) f_z(\ln \tilde{z}_{j,t}) f_z(\ln \tilde{z}_{j,t+dt} | \ln \tilde{z}_{j,t}) \\
= &\int_{\ln \tilde{z}_{\min}}^{\infty} d(\ln \tilde{z}_{j,t}) \int_{-\infty}^{\ln \tilde{z}_{\min}} d(\ln \tilde{z}_{j,t+dt}) \\
&\left(\tilde{z}^{\phi}_{\min} - \tilde{z}^{\phi}_{j,t+dt}\right) F_0 e^{-\lambda \ln \tilde{z}_{j,t}} \\
&\times \frac{1}{\sqrt{2\pi\sigma_z^2 dt}} e^{-\frac{\left(\ln \tilde{z}_{j,t+dt} - (\ln \tilde{z}_{j,t} + \tilde{\mu}_z dt)\right)^2}{2\sigma_z^2 dt}},
\end{aligned}$$

where $\tilde{\mu}_z \equiv \mu_z - g_z - \sigma_z^2/2 - m$, $f_z(\ln \tilde{z}_{j,t+dt} | \ln \tilde{z}_{j,t})$ is the distribution of $\ln \tilde{z}_{j,t+dt}$ conditional on $\ln \tilde{z}_{j,t}$ that follows a normal distribution, and $f_z(\ln \tilde{z}_{j,t})$ is the steady state firm size distribution.

Under the setup, taking the limit as dt approaches zero from above, (16) becomes

$$\begin{aligned}
\overline{\mathsf{E}}\left\{\tilde{z}^{\phi}_{j,t}\right\} m\phi &= \lim_{dt \to 0+} \frac{\overline{\mathsf{E}}\left\{\tilde{z}^{\phi}_{\min} - \tilde{z}^{\phi}_{j,t+dt}\middle|\tilde{z}_{j,t+dt} \leq \tilde{z}_{\min}\right\}}{dt} \\
&= \lim_{t' \to 0+} \frac{d\overline{\mathsf{E}}\left\{\tilde{z}^{\phi}_{\min} - \tilde{z}^{\phi}_{j,t+t'}\middle|\tilde{z}_{j,t+t'} \leq \tilde{z}_{\min}\right\}}{dt'}. \qquad (47)
\end{aligned}$$

$d\overline{\mathsf{E}}\left\{\tilde{z}_{\min}^{\phi}-\tilde{z}_{j,t+t'}^{\phi}\middle|\tilde{z}_{j,t+t'}\leq\tilde{z}_{\min}\right\}/dt'$ can be further calculated as follows:

$$
\begin{aligned}
&\frac{d\overline{\mathsf{E}}\left\{\tilde{z}_{\min}^{\phi}-\tilde{z}_{j,t+t'}^{\phi}\middle|\tilde{z}_{j,t+t'}\leq\tilde{z}_{\min}\right\}}{dt'}\\
&=\int_{\ln\tilde{z}_{\min}}^{\infty}d(\ln\tilde{z}_{j,t})\int_{-\infty}^{\ln\tilde{z}_{\min}}d(\ln\tilde{z}_{j,t+t'})\\
&\quad\frac{d}{dt'}\left(\left(\tilde{z}_{\min}^{\phi}-\tilde{z}_{j,t+t'}^{\phi}\right)F_0e^{-\lambda\ln\tilde{z}_{j,t}}\frac{1}{\sqrt{2\pi\sigma_z^2t'}}e^{-\frac{\left(\ln\tilde{z}_{j,t+t'}-(\ln\tilde{z}_{j,t}+\tilde{\mu}_zt')\right)^2}{2\sigma_z^2t'}}\right)\\
&=\frac{F_0e^{-(\lambda-\phi)\ln\tilde{z}_{\min}}\phi}{4\left(\lambda-\phi\right)}\\
&\quad\times\left(e^{\frac{1}{2}\lambda(2\tilde{\mu}_z+\lambda\sigma_z^2)t'}(2\tilde{\mu}_z+\lambda\sigma_z^2)\,\mathrm{Erfc}\left[\frac{(\tilde{\mu}_z+\lambda\sigma_z^2)\sqrt{t'}}{\sqrt{2}\sigma_z}\right]\right.\\
&\quad\left.-e^{\frac{1}{2}\phi\left(2\tilde{\mu}_z+\phi\sigma_z^2\right)t'}\left(2\tilde{\mu}_z+\phi\sigma_z^2\right)\mathrm{Erfc}\left[\frac{\left(\tilde{\mu}_z+\phi\sigma_z^2\right)\sqrt{t'}}{\sqrt{2}\sigma_z}\right]\right).
\end{aligned}
$$

By combining these results and taking the limit, we obtain

$$
\lim_{t'\to 0+}\frac{d\overline{\mathsf{E}}\left\{\tilde{z}_{\min}^{\phi}-\tilde{z}_{j,t+t'}^{\phi}\middle|\tilde{z}_{j,t+t'}\leq\tilde{z}_{\min}\right\}}{dt'}=\frac{1}{4}F_0e^{-(\lambda-\phi)\ln\tilde{z}_{\min}}\phi\sigma_z^2.
$$

Substituting this result into (47), we finally obtain

$$
m=(\lambda-\phi)\,\frac{\sigma_z^2}{4}.
$$

C Derivations of Households' Asset Distributions in the Steady State

This appendix shows the derivations of the households' asset distributions described in Section 5.

C.1 Derivation of the asset distribution of entrepreneurs

The discussion in Section 5.1 indicates that the probability density function of entrepreneurs at age t' with a detrended log wealth level of $\ln \tilde{a}_i$ is

$$f_e(\ln \tilde{a}_i|t') = \frac{1}{\sqrt{2\pi\sigma_{ae}^2 t'}} \exp\left(-\frac{(\ln \tilde{a}_i - (\ln \tilde{h} + (\mu_{ae} - g - \sigma_{ae}^2/2)t'))^2}{2\sigma_{ae}^2 t'}\right).$$

The probability density of entrepreneurs whose age is t' is

$$f_e(t') = \frac{(\nu + p_f)N}{L} \exp\left(-(\nu + p_f)t'\right).$$

By combining them, we can calculate the probability density function of the entrepreneurs' asset distribution, $f_e(\ln \tilde{a}_i)$, by

$$f_e(\ln \tilde{a}_i) = \int_0^\infty dt' \; f_e(t') f_e(\ln \tilde{a}_i|t').$$

To derive $f_e(\ln \tilde{a}_i)$ in Section 5.1, we apply the following formula to the above equation:

$$\int_0^\infty \exp(-at - b^2/t)/\sqrt{t}dt = \sqrt{\pi/a}\exp(-2b\sqrt{a}), \quad \text{for } a > 0, b > 0.$$

C.2 Derivation of the asset distribution of innate workers

The asset distribution of innate workers is calculated as follows:

$$\begin{aligned}
f_w(\ln \tilde{a}_i) &= \int_0^\infty dt' \; f_w(t') f_w(\ln \tilde{a}_i|t') \\
&= \int_0^\infty dt' \; \frac{\nu L - (\nu + p_f)N}{L} \exp(-\nu t') \cdot \mathbf{1}(\ln \tilde{a}_i = \ln \tilde{h} + (\mu_{a\ell} - g)t') \\
&= \int_{\ln \tilde{h}}^{\ln \tilde{h} + (\mu_{a\ell} - g)\infty} \frac{dt'}{d\ln \tilde{a}_i} d(\ln \tilde{a}_i) \; \cdot \frac{\nu L - (\nu + p_f)N}{L} \exp\left(-\frac{\nu}{\mu_{a\ell} - g}(\ln \tilde{a}_i - \ln \tilde{h})\right) \\
&\quad \times \mathbf{1}(\ln \tilde{a}_i = \ln \tilde{h} + (\mu_{a\ell} - g)t') \\
&= \begin{cases} \frac{\nu L - (\nu + p_f)N}{L} \frac{1}{|\mu_{a\ell} - g|} \exp\left(-\frac{\nu}{\mu_{a\ell} - g}(\ln \tilde{a}_i - \ln \tilde{h})\right) & \text{if } \frac{\ln \tilde{a}_i - \ln \tilde{h}}{\mu_{a\ell} - g} \geq 0, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

Note that $\mathbf{1}(\ln \tilde{a}_i = \ln \tilde{h} + (\mu_{a\ell} - g)t')$ is a unit function that takes 1 if $\ln \tilde{a}_i = \ln \tilde{h} + (\mu_{a\ell} - g)t')$ and 0 otherwise.

C.3 Derivation of the asset distribution of former entrepreneurs

The asset distribution of former entrepreneurs is derived as follows. Let $t'_m \equiv (\ln \tilde{a}_i - \ln \tilde{h})/(\mu_{a\ell} - g)$. First, we consider the case where $\mu_{a\ell} \geq g$. If $\ln \tilde{a}_i \geq \ln \tilde{h}$, then

$$\begin{aligned}
f_f(\ln \tilde{a}_i) =& \int_0^{t'_m} dt'\ p_f f_{e1}(\ln \tilde{a}_i - (\mu_{a\ell} - g)t') \times \exp(-\nu t') \\
&+ \int_{t'_m}^{\infty} dt'\ p_f f_{e2}(\ln \tilde{a}_i - (\mu_{a\ell} - g)t') \times \exp(-\nu t') \\
=& \left[\frac{-p_f}{\nu - \psi_1(\mu_{a\ell} - g)} f_{e1}(\ln \tilde{a}_i - (\mu_{a\ell} - g)t') \times \exp(-\nu t') \right]_0^{t'_m} \\
&+ \left[\frac{-p_f}{\nu + \psi_2(\mu_{a\ell} - g)} f_{e2}(\ln \tilde{a}_i - (\mu_{a\ell} - g)t') \times \exp(-\nu t') \right]_{t'_m}^{\infty} \\
=& \frac{p_f}{\nu - \psi_1(\mu_{a\ell} - g)} \left\{ -f_{e1}(\ln \tilde{a}_i - (\mu_{a\ell} - g)t'_m) \times \exp(-\nu t'_m) + f_{e1}(\ln \tilde{a}_i) \right\} \\
&+ \frac{p_f}{\nu + \psi_2(\mu_{a\ell} - g)} \left\{ -0 + f_{e2}(\ln \tilde{a}_i - (\mu_{a\ell} - g)t'_m) \times \exp(-\nu t'_m) \right\}.
\end{aligned}$$

By substituting the following relations into the above equation, $\ln \tilde{a}_i - (\mu_{a\ell} - g)t'_m = \ln \tilde{h}$, $f_{e1}(\ln \tilde{h}) = f_{e2}(\ln \tilde{h})$, and $t'_m = (\ln \tilde{a}_i - \ln \tilde{h})/(\mu_{a\ell} - g)$, we obtain,

$$\begin{aligned}
f_f(\ln \tilde{a}_i) =& \frac{p_f}{\nu - \psi_1(\mu_{a\ell} - g)} f_{e1}(\ln \tilde{a}_i) \\
&- \left(\frac{1}{\nu - \psi_1(\mu_{a\ell} - g)} - \frac{1}{\nu + \psi_2(\mu_{a\ell} - g)} \right) p_f f_{e1}(\ln \tilde{h}) \\
&\times \exp\left(-\frac{\nu}{\mu_{a\ell} - g} (\ln \tilde{a}_i - \ln \tilde{h}) \right).
\end{aligned}$$

If $\ln \tilde{a}_i < \ln \tilde{h}$,

$$\begin{aligned}
f_f(\ln \tilde{a}_i) =& \int_0^{\infty} dt' p_f f_{e2}(\ln \tilde{a}_i - (\mu_{a\ell} - g)t') \times \exp(-\nu t') \\
=& \frac{p_f}{\nu + \psi_2(\mu_{a\ell} - g)} f_{e2}(\ln \tilde{a}_i).
\end{aligned}$$

Next, we consider the case where $\mu_{a\ell} < g$. If $\ln \tilde{a}_i \geq \ln \tilde{h}$, then

$$
\begin{aligned}
f_f(\ln \tilde{a}_i) = & \int_0^\infty dt' p_f f_{e1}(\ln \tilde{a}_i - (\mu_{a\ell} - g)t') \times \exp(-\nu t') \\
= & \frac{p_f}{\nu - \psi_1(\mu_{a\ell} - g)} f_{e1}(\ln \tilde{a}_i).
\end{aligned}
$$

If $\ln \tilde{a}_i < \ln \tilde{h}$,

$$
\begin{aligned}
f_f(\ln \tilde{a}_i) = & \int_0^{t'_m} dt' p_f f_{e2}(\ln \tilde{a}_i - (\mu_{a\ell} - g)t') \times \exp(-\nu t') \\
& + \int_{t'_m}^\infty dt p_f f_{e1}(\ln \tilde{a}_i - (\mu_{a\ell} - g)t') \times \exp(-\nu t') \\
= & \frac{p_f}{\nu + \psi_2(\mu_{a\ell} - g)} f_{e2}(\ln \tilde{a}_i) \\
& - \left(\frac{1}{\nu + \psi_2(\mu_{a\ell} - g)} - \frac{1}{\nu - \psi_1(\mu_{a\ell} - g)} \right) p_f f_{e1}(\ln \tilde{h}) \\
& \times \exp\left(-\frac{\nu}{\mu_{a\ell} - g} (\ln \tilde{a}_i - \ln \tilde{h}) \right).
\end{aligned}
$$

D Derivation of the welfare analysis

In this appendix, we calculate the ex ante utilities of an entrepreneur and a worker in the steady state that were used in Section 6.5.3. We first derive the utility (value function) of a worker. By substituting (3) and (4) into (29) and rearranging, we obtain $H^w(\mathbf{S})$ in (32) in the steady state as follows:

$$
H^w(\mathbf{S}) = \frac{1}{\beta + \nu} \left[\ln(\beta + \nu) + \frac{r^f - \beta}{\beta + \nu} \right].
$$

By using this equation, the value function of a worker in the steady state, whose total asset is a_i, can be calculated by

$$
V^w(a_i, \mathbf{S}) = \frac{\ln a_i}{\beta + \nu} + H^w(\mathbf{S}).
$$

Next, using the above results, we derive the utility (value function) of an entrepreneur. From (29), we obtain $H^e(\mathbf{S})$ in (32) in the steady state as follows:

$$H^e(\mathbf{S}) = \frac{1}{\beta + \nu + p_f}\left[p^f H^w(\mathbf{S}) + \ln(\beta+\nu) + \frac{r^f - \beta + (\mu^q - r^f)x_e/2}{\beta+\nu}\right].$$

The value function of an entrepreneur in the steady state, whose total asset is a_i, can be calculated by

$$V^e(a_i, \mathbf{S}) = \frac{\ln a_i}{\beta+\nu} + H^e(\mathbf{S}).$$

Section 6.5.3 calculates the detrended utility level of an entrepreneur and an innate worker that is defined by

$$\widetilde{V}^i(\tilde{h}, \mathbf{S}) \equiv V^i(h, \mathbf{S}) - gt = \frac{\ln \tilde{h}}{\beta+\nu} + H^i(\mathbf{S}).$$

	Pre-1975	Post-1975
Ordinary income tax, τ^{ord}	0.75	0.40
Corporate income tax, τ^{corp}	0.50	0.35
Capital gain tax, τ^{cap}	0.25	0.25
τ^{e}	0.75	0.40
τ^{f}	0.63	0.51

Table 1: Tax rates

Notes: The figures in the upper half of the table are calibrated from the top statutory marginal federal tax rates in Figure 1 that is taken from Saez et al. (2012). The tax rate on risky stocks, τ^{e}, is set to be equal to τ^{ord}. The tax rate on risk-free assets, τ^{f}, is calculated by $1-(1-\tau^{\text{cap}})(1-\tau^{\text{corp}})$.

β	Discount rate	0.04
ν	Prob. of death	1/50
α	Capital share	1/3
δ	Depreciation rate	0.1
g	Steady state growth rate	0.02
ρ	Elasticity of substitution	0.7
p_f	Prob. of entrepreneur's quitting	1/20
$\ell_{\min}$	Min. level of employment	1
L	Mass of population	1.0
N	Mass of entrepreneurs	0.05

		Case A	Case B
$\phi\sigma_z$	Firm-level vol. of employment	0.25	0.45
ι	Transaction costs of fin. intermed.	0.215	0.243

Table 2: Calibrated parameters

Notes: The figures of the firm-level volatility of employment are taken from Figure 2.6 of Davis et al. (2007). Case A corresponds to the case where firm-level volatility is equal to that of publicly traded firms in the data. Case B corresponds to the case where firm-level volatility is equal to that of both publicly traded and privately held firms in the data.

	Case A	Case B
Pre-1975	4.7%	6.2%
Post-1975	1.9%	6.3%

	Data
1980	4.9%
2006	8.3%

Table 3: Size of the financial sector

Notes: The left table shows the model's predictions on the size of the financial sector over GDP $\iota\left(1-\frac{\tilde{A}_{e,t}x_t}{\tilde{Q}_t}\right)\tilde{D}_t/\tilde{Y}_t$ at the pre- and post-1975 steady states under the parameter values in Cases A and B. The right table shows the share of the financial sector in GDP in the U.S. in 1980 and 2006. These data are taken from Greenwood and Scharfstein (2013).

Case A		
	$\widetilde{V}^e(\tilde{h}, \mathbf{S})$	$\widetilde{V}^w(\tilde{h}, \mathbf{S})$
Pre-1975	36.27	35.03
Post-1975	36.55	32.84

Case B		
	$\widetilde{V}^e(\tilde{h}, \mathbf{S})$	$\widetilde{V}^w(\tilde{h}, \mathbf{S})$
Pre-1975	36.23	34.98
Post-1975	35.63	33.12

Table 4: Welfare analysis

Notes: The table calculates the detrended initial utility level of an entrepreneur and an innate worker at the pre- and post-1975 steady states. The detrended initial utility level is defined by $\widetilde{V}^i(\tilde{h}, \mathbf{S}) \equiv V^i(h, \mathbf{S}) - gt$. The left table presents these calculations for Case A, whereas the right table presents them for Case B.

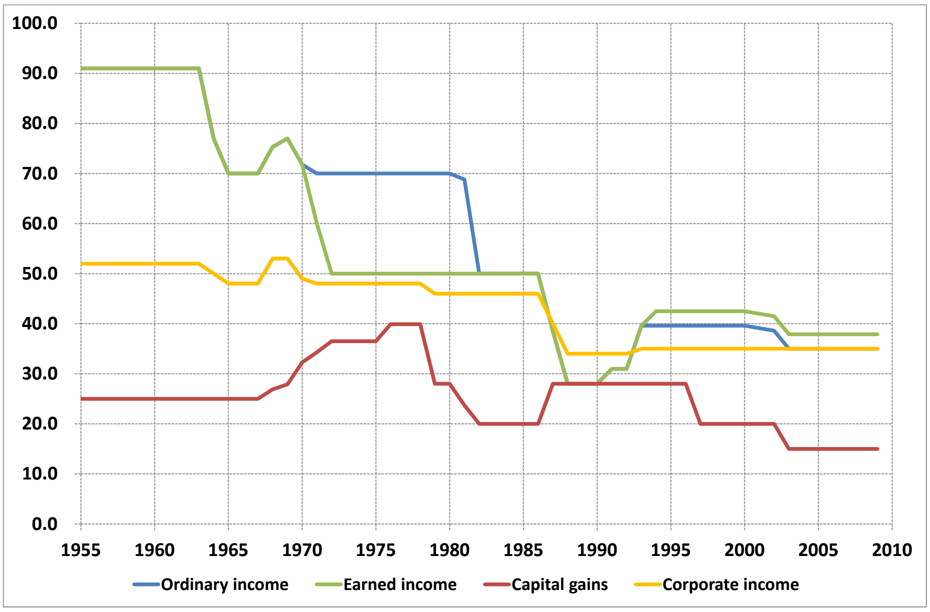


Figure 1: Federal tax rates

Note: The data are taken from Table A1 of Saez et al. (2012).

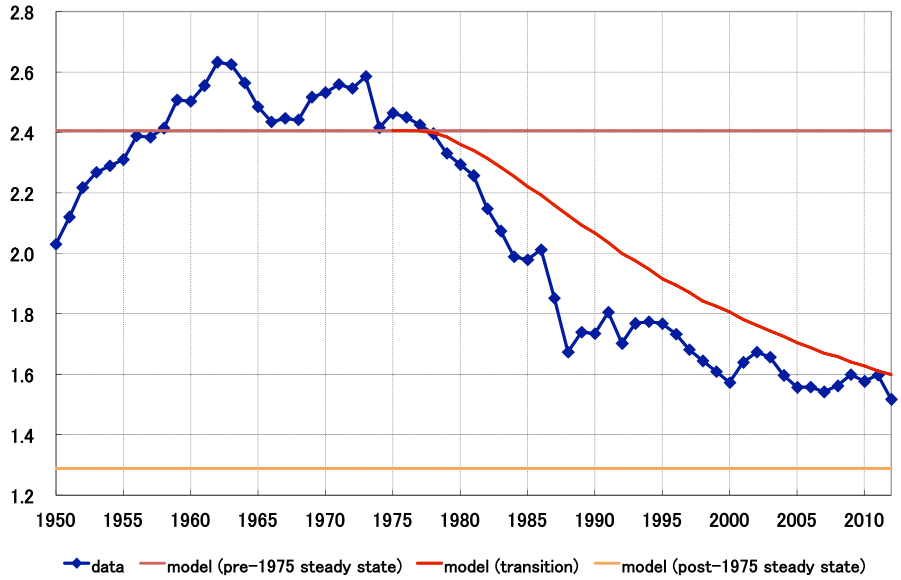


Figure 2: Pareto exponent: Case A

Note: Data are taken from Alvaredo et al. (2013).

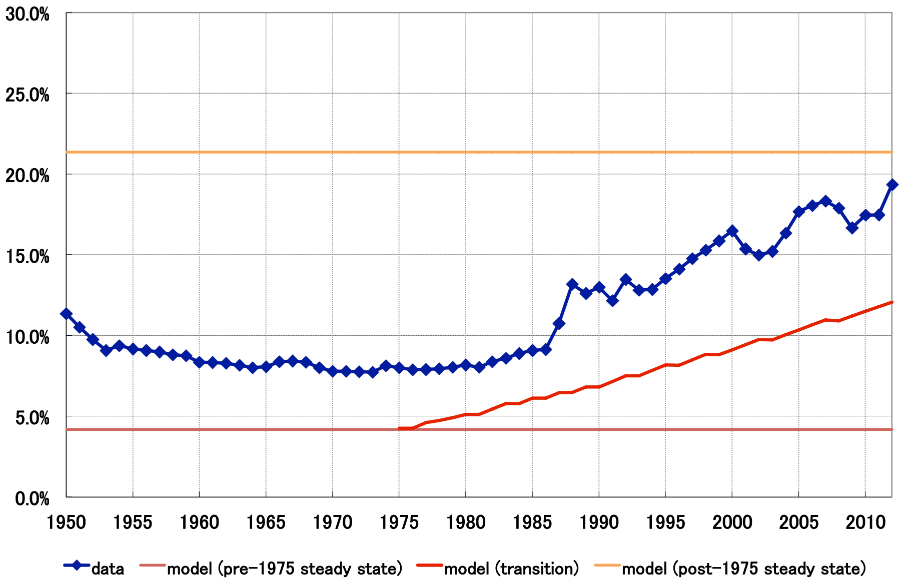


Figure 3: Top 1% share of income: Case A

Note: Data are taken from Alvaredo et al. (2013).

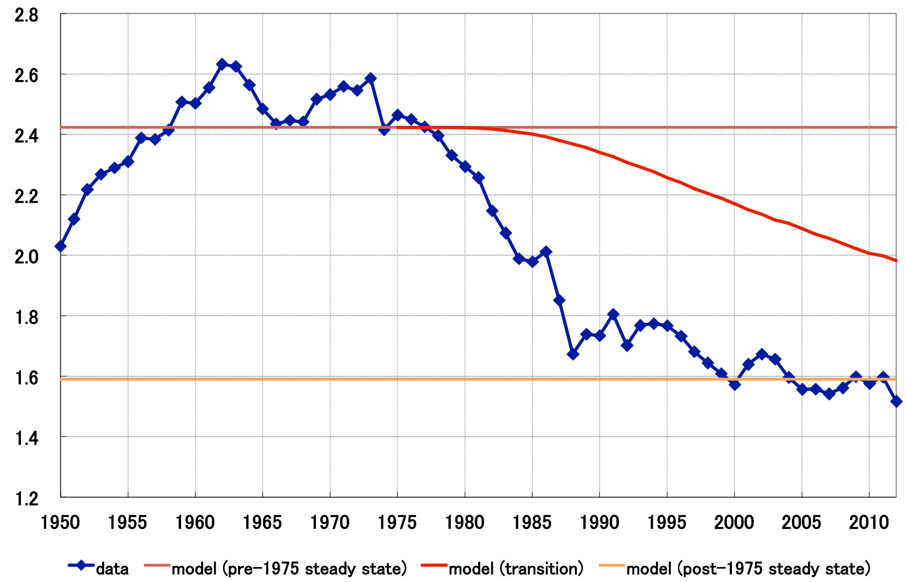


Figure 4: Pareto exponent: Case B

Note: Data are taken from Alvaredo et al. (2013).

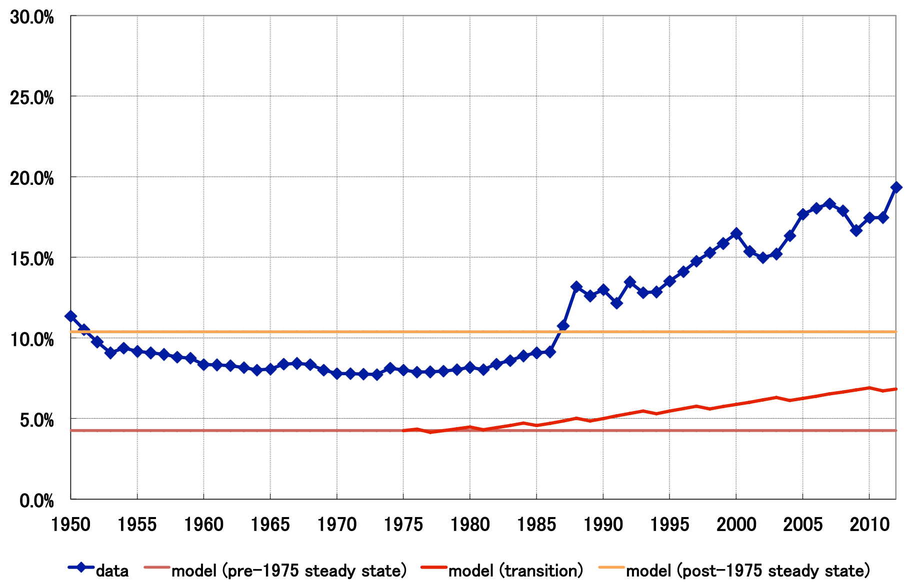
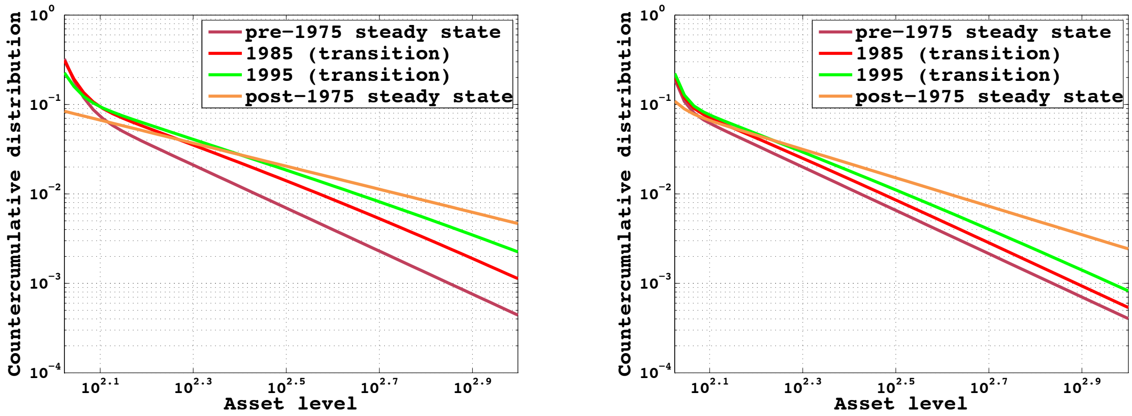


Figure 5: Top 1% share of income: Case B

Note: Data are taken from Alvaredo et al. (2013).



(a) Case A

(b) Case B

Figure 6: Household's asset distributions

Notes: The figures plot the countercumulative distributions of the household's detrended asset under the pre- and post-1975 steady states as well as the transition paths. For example, "1985 (transition)" indicates the wealth distribution in 1985 under the model's transition path. The left figure presents the distributions for Case A, whereas the right figure presents them for Case B.

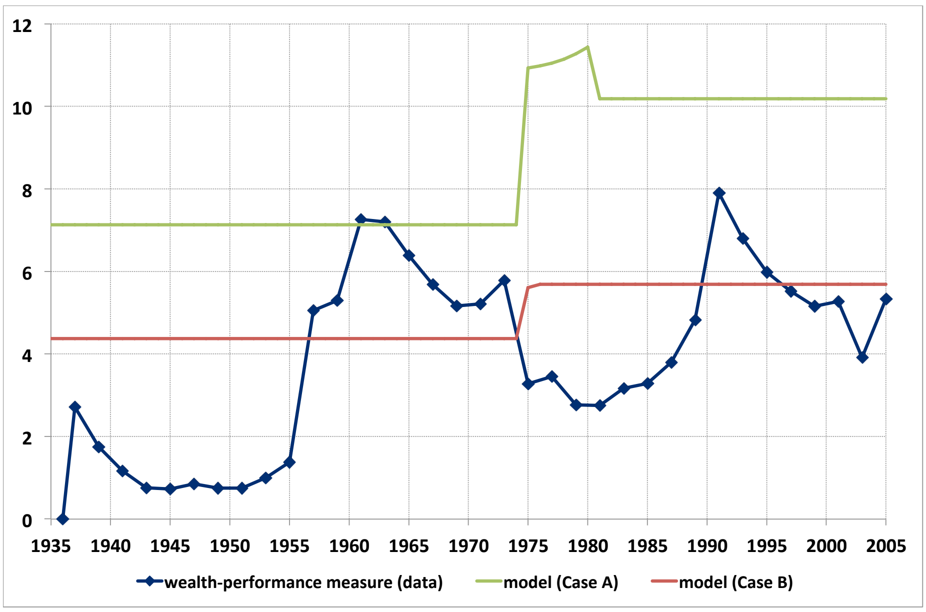


Figure 7: Wealth–performance measure

Notes: For the definition of the wealth–performance measure, see (28). The data are calculated by dividing "dollar change in wealth for a 1% increase in the firm's rate of return" by "total compensation," both of which are estimated in Frydman and Saks (2010). These data correspond to the median values of the 50 largest firms.